Journal of
Experimental Psychology

VOL. 27, NO. 4 OCTOBER, 1940

# THE ROLE OF HEAD MOVEMENTS AND VESTIBULAR AND VISUAL CUES IN SOUND LOCALIZATION [1]

BY HANS WALLACH

*Swarthmore College*

In a previous paper [2] the writer has demonstrated that a distinct localization of sound exists for directions which do not fall into the horizontal plane but lie above or below at varying elevations, in other words that a discrimination of directions with respect to above and below and front and back is possible as well as discrimination with respect to right and left which has been studied for many years. It was found that only a head movement during the presentation of the sound affords an adequate discrimination of sound direction in the dimension of above and below and thus makes localization complete. This is probably the reason why sound localization with respect to above and below has not been demonstrated in the laboratory at an earlier time. The paper referred to gives an analysis of the manner in which a complete perception of a sound direction is achieved with the help of a head movement. The binaural cues on which sound localiza-

[1] The writer wishes to express his gratitude to Prof. W. Koehler and to Dr. Mary Henle for their help in preparing this manuscript.

[2] Ueber die Wahrnehmung der Schallrichtung. *Psychol. Forsch.*, 22, pp. 238–266 (I). A short English report on the same work has been published in the *J. Acoust. Soc. Amer.*, 1939, pp. 270–74 (II). The two papers will be referred to as I and II respectively. In the present paper the writer has attempted to present the material in such a way that it can be understood without knowledge of the previous papers. While paper II lacks theoretical discussion it may be useful for a quick reference to previous experimental results.

tion is primarily based do not suffice to characterize a sound direction completely. Yet, during a head movement the binaural cues as produced by the sound direction are altered, and the particular form of this change can in each case strictly determine the given sound direction. This is thoroughly discussed in the article mentioned. Experiments are reported in which perceived sound directions were synthetically produced in accordance with the head movement principle, which was thus verified. It is obvious that the change which the primary factors undergo due to the head movement can characterize a sound direction only if the exact kinematic properties of the particular head movement are taken into account. Two sets of sensory data enter into the perceptual process of localization, (1) the changing binaural cues and (2) the data representing the changing position of the head. It is the latter with which this paper is concerned. The manner in which they are secured is the object of this investigation.

## I. The Role of Head Movement in Auditory Localization

It has been shown that the binaural cues for sound localization, time difference and difference of intensity, convey the angular distance of the given sound direction from the axis of the ears. Thus, they only determine how far from the median plane, on the left or on the right side, the given source of sound is located. Whether it lies in front or in the rear, above or below the horizontal plane, remains undetermined, and the same is true of the amount of its elevation. The angular distance of the given sound direction from the aural axis which is actually determined by the binaural cues has been called the lateral angle ($\psi$), and it is treated in this paper as if it were a directly given sensory datum.[3] This lateral angle can be counted from either pole of the aural axis.

[3] This is the terminology employed in paper II. In paper I, the term lateral angle refers to the complement of the angle $\psi$ and is called $\varphi$. $\varphi$ is thus the angular distance of the sound direction from the median plane and was chosen to represent the lateral angle for historical reasons. The motive for changing to $\psi$ was the greater ease with which the angle $\psi$ can be visualized.

If it is 90°, the sound direction lies in the median plane of the head. It follows that *the binaural cues determine merely a range of directions, any of which would, if actually presented, produce the same binaural stimulation.* It is significant that such a range of possible directions for a given sound is never perceived. We hear a sound that appears for the most part in one definite direction. It has been shown that this is due to head movements during the perception of the sound. A motion of the head will, in most cases, alter the position of the aural axis and at the same time change the angle between the latter and the given sound direction. It will be seen shortly that this change of the lateral angle can define the direction of the source of sound; and actually the perception of the proper direction is achieved through a head movement affording such a change of the lateral angle.

In order to show that the change of the lateral angle with the head movement can define a direction, we shall consider a number of particular cases. We shall at first assume that the head is turned about a vertical axis, so that the displacement of the aural axis occurs in the horizontal plane. If the given direction lies in the same plane, the angle between this direction and the aural axis (the lateral angle) changes by the amount of the displacement of the aural axis. This is obviously no longer true when the given direction is not in the horizontal plane, but is above or below. While in the case of the horizontal direction the lateral angle is measured within the horizontal plane, in the case of an elevated direction the angle between this direction and the aural axis extends in an oblique plane. It can easily be seen that this angle is affected to a lesser degree by a shift of the aural axis in the horizontal plane than is the lateral angle in the case of a horizontal direction. If the reader finds himself unable to visualize these spatial relations, the following consideration will lead to the same conclusion.

Take the case in which the direction is exactly vertical. The direction then exactly coincides with the axis of the head movements, and the displacement of the aural axis will not alter the lateral angle at all. It remains 90° throughout the head

movement. This is the extreme case. For all other directions the lateral angle will be changed by a displacement of the aural axis. A direction may be given 60° above the horizontal plane, and, before the head movement starts, exactly in front. At this stage, the lateral angle amounts to 90°. When a movement of the head by 90° brings the direction in the lateral position, the lateral angle will amount to 60°, the aural axis still lying in the horizontal plane with the direction 60° above it. A 90° shift of the aural axis thus brings about a 30° change of the lateral angle. For a horizontal direction, on the other hand, a like displacement brings about a 90° change of the lateral angle, as we have seen, while for the vertical direction the lateral angle does not change at all. These three cases are sufficient to suggest what can be learned directly from a visualization of the spatial relations, namely, that the amount by which the lateral angle changes varies with the elevation of the direction. For a given head movement this change is maximal when the direction lies within the horizontal plane, and decreases as the direction approaches the vertical.

Thus the angular distance of the direction from the horizontal plane (angle of elevation) varies with the rate of change of the lateral angle and is determined by it. For a given lateral angle, there are however in most cases four directions, which have the same angular distance from the horizontal plane: in front and in back, above and below. Of these, the two in front are distinguished from the two in back by another feature of the change of lateral angle, the direction, viz., the sign, of the change. When the head is turned to the left a direction in front will shift toward the right side of the head. For the same head movement, a direction in back will shift toward the left side of the head. That is to say, for a direction in front the lateral angle decreases toward the right, for a direction in back it decreases toward the left, and the two directions may thus be distinguished, although both may have the same elevation, and consequently the lateral angle will change for both at the same rate. Only one ambiguity remains: two directions in symmetrical posi-

tion with respect to the horizontal plane, one above and the other below, are so far not distinguished.

These considerations concerning the rate and the direction of the change of lateral angle are independent of the particular spatial orientation of the head movement. The angle of elevation does not, of course, refer to the horizontal plane as such, but to the plane in which the aural axis is displaced by the head movement, which in the particular case we chose coincided with the horizontal plane. Our considerations apply as well to such a case as a movement about a horizontal axis, that is, a tilting of the head from side to side. Here the aural axis shifts in a vertical plane, the 'angle of elevation' extends toward the front or the rear, and above and below now play the same rôle as front and back did in the case in which the head was turned. They are distinguished from one another through the direction of the change which the lateral angle undergoes due to the tilting motion.

When, upon a movement of the head about a vertical axis, two directions remain which are both consistent with the given change of lateral angle, one above and the other below the horizontal plane, a subsequent tilting of the head, or any motion which contains the tilting as a component, can distinguish between the two possibilities. In fact, natural head movements are rarely accurate revolutions about a constant axis; they must rather be described as revolutions with varying axis. This does not interfere with the qualification of the head movement for procuring a change of lateral angle which can determine a sound direction to the extent indicated above. But the displacement of the axis which occurs during the head movement probably suffices to remove the ambiguity which would result from an accurate revolution about a constant axis.[4]

We have found a number of different movements of the head to be effective in sound localization. The most frequent natural head movement is a turning of the head upon which a tilting to the side is gradually superimposed as the motion approaches the end of the excursion. In the syn-

[4] For a more detailed discussion of this point *cf.* 1 §7.

thetic production of sound directions which were previously reported, a revolution about a vertical axis, in which all components of tilting to the side had to be strictly excluded, was successfully used. In another group of these experiments the head movement was even more artificial; it consisted of a tilting of the head from side to side. With such unnatural movements most accurate localizations were achieved. We are probably justified in saying that any movement of the head which involves an angular displacement of the aural axis can be effective in sound localization. That a head movement which does not involve such a displacement of the aural axis must be ineffective is evident. No change of lateral angle can result from it. Such a head movement is a revolution about a horizontal axis extending from the left to the right, as it occurs in a straight nodding. It is, of course, also ineffective as a component in a kinematically more complicated natural head movement. Briefly, any head movement may be effective to the extent to which it contains as a component a revolution about an axis which lies in the median plane of the head.

The tests which will be reported in the next section are performed with synthetically produced sound directions. The manner in which sound directions can be synthetically produced has been thoroughly discussed in a previous paper; thus only a general outline will be presented in the following.

In experiments on synthetic production, the binaural stimulation which a given objective sound direction would have produced is brought about without actual presentation of this sound direction. We have seen that for a given position of the head the binaural stimulation which is produced by a given sound corresponds to quite a number of directions. All directions which have the same lateral angle produce the same binaural stimulation. Thus, so far as binaural cues are concerned, any one of these directions can be substituted for any other one. Where the lateral angle of a given direction changes in a characteristic way during a head movement, this direction can be replaced by a series of other directions which together present the same changing lateral angle. For

every position through which the aural axis passes the lateral angle can be ascertained, and an equivalent direction can be substituted for the given direction. In practice one need not even consider an infinite number of such positions because a differential threshold exists in discrimination of the lateral position of a sound. Of two positions of the head for which the lateral angle is merely subliminally different, only one need be taken into account. We find it sufficient to consider separate positions which are as much as 3° apart. For the synthetical production of a certain sound direction, a definite head movement must be selected. Separate positions through which the aural axis passes during this head movement are chosen 3° apart from each other, and for each of these positions the lateral angle of the direction which is to be produced is ascertained. Thereupon, for each of these positions a direction can be selected among those which have the specific lateral angle ascertained for this position. In this way a series of directions is obtained which are equivalent to the direction which is to be produced, if each is presented at the moment when the head in its movement passes through the position to which it belongs. For the presentation of each of these different directions at the proper moment, the head of the observer is attached to a switch with 20 contact points. Each of these contact points is connected with a loudspeaker. In a particular experiment the switch permits only that movement of the head which is chosen in planning the experiment. It is always a revolution about a constant axis lying in the median plane of the head. During such a movement a contact spring slides over the contact points one after another, thus connecting each loudspeaker in turn. With every 3° displacement of the aural axis the center of another contact point is passed. The position of head and aural axis at the moment when the center of each contact point is passed can easily be ascertained, and the corresponding loudspeaker is placed in the direction which has previously been chosen to present the proper lateral angle for this particular phase of the head movement. When all the loudspeakers are arranged in this manner, the apparatus achieves precisely what

is necessary for the synthetic production of the sound direction. While the head is moved, directions are presented which are equivalent to the direction to be produced. Together they present the same sequence of lateral angles which the synthetic sound direction would bring forth if it were actually given. In such experiments the physically given sound directions differ widely from the synthetic directions which are perceived.

In previously reported experiments the twenty loudspeakers which presented the sound were arranged in front of the observer, while the synthetic direction was, for instance, in back of or above the observer's head. In another experimental arrangement the observer perceived a sound directly in front, while all the actually presented directions lay above his head, distributed from left to right. Localization functioning on a basis other than the head movement principle should lead to the perception of sound in the directions actually presented or in their general neighborhood. If nevertheless the synthetic direction is perceived, one can be sure that it is solely on the basis of the head movement principle.

## II. Passive Displacement of the Aural Axis

So far we have consistently used the term head movement. Yet, according to the discussion in Section I any change of position of the head which involves an angular displacement of the aural axis, no matter how it is brought about, must suffice to procure a characteristic change of lateral angle. In the abstract, it should not matter whether an observer turns his head actively, or turns on his heel, or swings passively on a revolving chair, if the same displacement of the aural axis results. We have already pointed out, however, that two sets of sensory data enter into the perceptual process of sound localization, the changing binaural cues which represent the change of the lateral angle and the data which characterize the displacement of the aural axis. Without the latter, the change of lateral angle has no significance for sound localization. The nature of the stimulation which represents the displacement of the head will vary with the manner in which

the displacement is achieved, and therefore this manner may be very significant. In fact, any stimulation which is capable of coöperating with the binaural cues in sound localization should be investigated. For this reason the various ways in which the head can be displaced will now be examined.

Three kinds of sensory data may represent a displacement of the head: proprioceptive stimulation from the muscles engaged in active motion, stimulation of the eyes, and stimulation of the vestibular apparatus. Usually all three are present, but in experimentation one or another may be excluded, and thus the significance of each kind may be examined. However, one of them, the vestibular stimulation, cannot be eliminated where spatial displacement of the head is involved; it will be always present. It should therefore be investigated first.

In order to eliminate the optical cues for his motion and the proprioceptive stimulation which would result from active bodily movements, the observer was blindfolded and placed on a revolving chair on which he could be turned about a vertical axis. A back rest and a foot rest as well as a rest for the chin served to decrease the slight stresses which arise when the acceleration is conveyed through the body from the parts in contact with the chair. Except perhaps for stimulation resulting from these stresses, the observer was made aware of the chair's and his own movements only through the function of the vestibular apparatus. The sounds which were to be localized under these circumstances were presented in exactly the same way as in the experiments on synthetic production. This was necessary because of the existence of a secondary factor in sound localization, the effect of the pinnæ,[5] which has no connection with the head movement principle, and which might possibly have affected the results. By presenting a synthetic sound direction, a differentiation between the effect of the head movement and the effect of this secondary factor becomes possible.

In the experiments we shall report here the displacement of the head was a revolution about a vertical axis. The loud-

[5] *Cf.* I §18 and II p. 273.

speakers were arranged in a single row in the horizontal plane, *i.e.*, the same plane in which the aural axis was displaced. The first test made with passive displacement and with blindfold was the synthetic production of a horizontal direction in back, with the row of loudspeakers in front. For this experiment the loudspeakers had to be arranged at distances of 6° from each other, the angles measured from a point which in the actual experiment coincided with the center of the observer's head.

It can easily be shown that this arrangement is suitable for the synthetic production of the desired direction: For a certain position of the head, the direction to be produced may lie straight in back. To this position of the head a position of the switch may correspond in which the contact spring touches exactly the center of a contact point. The loudspeaker which is connected with this contact point is placed straight in front of the observer. Thus its direction is equivalent to the desired direction, for in the case of both directions the lateral angle is 90°. Now we assume that the observer turns his head 3° to the right. This displacement turns the switch by one contact, and at the same time brings the desired direction in back closer to the right pole of the aural axis. The lateral angle of the desired direction decreases by 3° and is 87° for the resulting position of the head. The loudspeaker which is connected with this new contact point must be placed in a direction which has a lateral angle also of 87°. When this is done, its position is 6° to the right of the previously placed loudspeaker, that is, 3° closer to the right pole of the aural axis on account of the change of lateral angle, plus another 3° because the aural axis itself is displaced to the right by 3° with respect to the position for which the first loudspeaker was placed. The center of the following contact point is reached after another 3° of displacement of the aural axis; the lateral angle of the desired direction in back decreases again by 3°, and again the corresponding loudspeaker must be fixed in a position 6° to the right of the one placed just before, that is, 3° to account for the changed position of the aural axis and another 3° to account for the decrease in

the lateral angle. The same procedure is, of course, applicable to the placing of all the following loudspeakers in the row and can also be applied to those which lie to the left of the loudspeaker placed first. Each is 6° distant from the next.

In an actual experiment when the observer's head is turned to the right by a certain angle, the sound will shift to the right by twice that angle, appearing in one loudspeaker after another, and when the head is turned to the left, the sound will be displaced to the left in the same manner. Thus, as the head is turned back and forth the sound slides through the row of loudspeakers at twice the rate of the rotation of the head. To stress again the essential point: this sliding sound will present to the observer the same sequence of lateral angles which characterizes the desired direction in back, no matter how quick or how wide the excursion of the head may be, provided that it occurs within the range of the switch. With 20 contact points 3° apart, the switch allows a maximum displacement of the head of 60°, and stops on the switch confine the head movement to this range. The 20 loudspeakers 6° apart cover an angle of 120°, and this is the corresponding maximum displacement of the sound.

Before the observer was seated, the position of the revolving chair was carefully adjusted so that its axis of revolution was in line with the axle of the switch which was, of course, in these experiments, also in a vertical position. After the observer had been placed on the chair, he adjusted his posture until his head could be fastened to the switch; for the switch had to remain in its position relative to the chair, and only its height above the floor could be altered. When the head was fastened, the chin rest was moved up and was fixed in the proper position. Then the observer was blindfolded. In all experiments reported in this paper, orchestra or piano music from victrola records was presented. When the music was turned on, it issued from one of the loudspeakers. From which one it came depended upon the accidental position of head and switch at the moment. The experimenter began to turn chair and observer back and forth, and the physical sound underwent the corresponding shifts. After a few ex-

cursions, interrupted by short pauses, the music was interrupted, and the observer was asked from where he had heard it.

All five observers who took part in this experiment heard the sound in the desired direction in back. When the experiment was varied in such a way as to bring the loudspeakers in back of the observer, the result was again positive for all observers: the sound appeared in front. As a further check, the switch was so constructed that the mechanical connection between the attachment of the head, on the one hand, and the switch proper, on the other hand, could be interrupted. In this case the contact spring remained at rest and one particular loudspeaker was constantly connected, while the observer was turned back and forth as before. When this was done during the presentation, the place of the perceived sound changed immediately; now it appeared in the direction of the sounding loudspeaker.

In the synthetic production, the sound was not only perceived on the reverse side in all cases, but in 9 of the 10 trials its direction was also exactly horizontal. Only in one case did the sound appear 25° above the horizontal plane. It will be shown below that under the conditions in question this direction could not possibly have been distinguished from a horizontal direction, and that it likewise represents a satisfactory result.

The positive results of these tests show clearly that sound localization based on the head movement principle is possible, even if the observer is passively turned and blindfolded. Yet they do not enable us to make an estimate of the accuracy with which the localization is achieved under these conditions. Firstly, localization within the horizontal plane, the only one we have as yet demonstrated, seems to be generally favored. If, for some reason, the data given for the localization of an elevated direction are inadequate, the sound is usually heard in a horizontal direction. Secondly, under our experimental conditions, the displacement of the aural axis occurs strictly in the horizontal plane. It has previously been pointed out that sound directions near the plane in which the aural axis

is displaced (the equatorial plane of the head movement) are poorly defined. The reason for this fact is purely geometrical. In order to ascertain numerically how much the lateral angle of an elevated direction changes with a given head movement, one makes use of the following formula

$$\sin (90° - \psi) = \sin \beta \cdot \cos \vartheta,^{6}$$

where $\psi$ stands for the lateral angle (the angle between the sound direction and the aural axis), $\vartheta$ for the angle of elevation of the given direction above the plane in which the aural axis is displaced, and $\beta$ for the angle by which the aural axis is displaced within this plane. According to this expression, the rate at which $\psi$ changes with changing $\beta$ is largely dependent upon the cosine function of the angle $\vartheta$. Now the cosine function shows little change for the angles between 0° and, say, 30°. Applied to the present case this means that for directions of smaller elevation the rate of change of the lateral angle is practically the same over a considerable range of directions, and a direction which is determined by such a rate of change will be poorly defined in spatial terms. This, incidentally, furnished the explanation for our accepting a 25° elevated direction as a satisfactory result in the experiment under discussion. The rate of change for this direction is not different enough from the one which characterizes a horizontal direction to afford a discrimination. It may be added that under ordinary circumstances these conditions do not constitute a real difficulty for sound localization. As noted above, natural head movements are revolutions with varying axis, and with an approach of the axis toward the given direction conditions for a sharp determination of this direction improve rapidly.

In order to determine how accurately a blindfolded and passively moved observer can localize, and especially how his achievements compare with results under normal conditions, a direction was synthetically produced which had an elevation of 60° above the horizontal plane—in back when the loudspeakers were in front, or vice versa.

[6] *Cf.* 1 p. 245, where $\varphi = 90° - \psi$.

The arrangement of the loudspeakers followed the same general procedure as that employed in the first experiment. As the desired 60° direction, one in back of the observer's seat was chosen. That contact point on the switch was ascertained which was connected when this direction fell into the median plane of the observer's head; and the corresponding loudspeaker was placed straight in front of the observer. A displacement of the aural axis by 3° would change the position of the switch so that the next contact would be closed. At the same time the aural axis would come somewhat nearer the desired direction, and the lateral angle would decrease. Yet this decrease would not amount to a full 3° as in the first experiment. An elevated direction, we have seen, corresponds to a change of lateral angle which is smaller than the displacement of the aural axis by which it is caused. Just how great this change will be can be computed from the formula given above. We let $\beta$ equal 3° and substitute for $\vartheta$ the value 60°, that is the elevation of the desired direction above the horizontal plane, the plane in which the aural axis is displaced. Thus we obtain for $\psi$ the value 88° 30′. Once we know the value of the lateral angle, the loudspeaker corresponding to the contact point in question can easily be placed. The loudspeaker which was placed first had a median position; its lateral angle was 90°. If the displacement of the aural axis is clockwise, the new loudspeaker must be placed to the right of the first one, that is, closer to the right pole of the aural axis; for a clockwise displacement brings the desired direction in back nearer to this pole. In order to determine its exact place we have to take into account the fact that, because of the displacement, the first loudspeaker is no longer in a median position. It now lies 3° to the left of the median plane, that is 93° from the right pole of the aural axis. Thus the new loudspeaker must be placed 4° 30′ to the right of the one placed first in order to make its lateral angle 88° 30′. In other words, of the 4° 30′ angular distance between the two loudspeakers, 3° account for the different positions of the aural axis when one or the other speaker is connected, and 1° 30′ account for the change

of lateral angle which takes place when the head passes from one position to the other.

The position of the next loudspeaker may be found in the same way. $\beta$ has here the value 6°, while $\vartheta$, of course, remains 60°, and the result obtained for $\psi$ is 87°. For the corresponding position of the head the first loudspeaker lies 6° to the left of the median plane, while the new one must be placed $90° - \psi = 3°$ to the right of the median plane, *i.e.*, $\beta + 90° - \psi = 9°$ to the right of the loudspeaker placed first. This loudspeaker is the starting point with reference to which the other loudspeakers are placed. In this manner all the loudspeakers on the right of the first are arranged. Those on the left side are to be arranged in the same way, with the qualification that here all angles refer to the left pole of the aural axis. When the head of the observer is fastened to the switch he must face the first loudspeaker directly, while the position of the switch must be such that this loudspeaker is connected.

In the actual test the procedure was exactly the same as in the first experiment. The experimenter turned chair and observer swiftly back and forth so that the velocity of the displacement approximated the speed of active head movements. Here again the observer could face the loudspeakers and have the desired direction in back, or vice versa. It seems that the situation in which the desired direction is to appear in front is more favorable, because directions in front are psychologically better defined than directions in back. Thus, most of the tests were done in this manner.

With elevated directions there is always the problem of how the observer is to indicate the direction of the sound. He can give an estimate of the elevation in degrees or he can point in the direction of the sound. Neither method is very accurate, and constant errors probably occur in both. The observer was allowed to choose between the two methods and encouraged in the use of both at the same time. Where both data are available they are indicated in the table. With one exception the 15 observers had only one trial.

TABLE

| Subject | | Blindfolded | Not Blindfolded |
|---|---|---|---|
| I | *b* | 60° *p*, 60° *est* | same |
| II | *b* | 60° *est* | 60° *est* |
| III | *b* | 55° *p*, 50° *est* | same |
| IV | *b* | 60°–80° *est* | 52° *p* |
| V | *f* | 65° *p*, 80° *est* | same |
| V | *b* | 55° *p* | same |
| VI | *b* | 50° *p*, 60° *est* | 60° *p*, 60° *est* |
| VII | *f* | 45°–50° *est* | same |
| VIII | *b* | 45° *p* | 65° *p* |
| IX | *f* | 40° *p*, 40° *est* | 60° *p* |
| X | *b* | 35° *p* | 41° *p* |
| XI | *b* | 30° *est* | 70° *est* |
| XII | *f* | 30° *est* | 50°–60° *est* |
| XIII | *b* | 20° *p* | 70° *p* |
| XIV | *b* | 10°–20° *est* | 53° *p* |
| XV | *b* | 10° *p* | 64° *p* |

Desired direction = 60°.
*b* = loudspeakers in back.
*f* = loudspeakers in front.
*p* = subject points in the direction of sound image.
*est* = subject estimates its elevation.

The results are given in the first column of the table. The observers are listed in the order of their success in this experiment. The results of the first 6 are satisfactory, and the remaining judgments are too low.

With each observer a variation of this experiment was performed. The blindfold was removed from the observer's eyes, the experimenter gave the impression of changing the apparatus so as to suggest quite a different experiment, and then proceeded just as he had done before. The results are listed in the second column of the table. Without the blindfold, 13 of the 15 observers showed quite satisfactory results. This means that 7 observers showed a definite improvement over their performance when blindfolded. The two who did not give satisfactory results (VII and X) were tested again with the chin rest removed so that active head movements became possible. The results showed no change of the direction they perceived. Apparently these subjects belonged to that large group who generally underestimate the elevation of any steep direction. The constancy in the results of these observers makes this interpretation plausible.

Two facts emerge from these results: Firstly, a sizable number of observers is able to localize sound adequately with passive displacement of the head and with exclusion of sight, that is, largely on the basis of vestibular stimulation. Secondly, all observers when passively moved are able to localize if vision is *not* excluded at least as well as they would do with active head movements. Under the given experimental conditions the first fact is quite remarkable. A revolution of the head strictly about a vertical axis does not entail a change of the direction of the gravitational force relative to the vestibular apparatus; only forces of acceleration give rise to stimulation of the vestibular system. This means that no direct cues for the position of the head at a given moment are obtained in the vestibular system. Even the velocity of the displacement of the head can only indirectly be derived from the stimulation; for only acceleration and not velocity as such can stimulate the vestibular apparatus. In whatever manner the displacement of the head is functionally represented in the process of sound localization, whether as a sequence of positions or as a rate of change,[7] the original sensory data have to undergo a transformation. These circumstances together with the fact that here the vestibular data are to be evaluated quantitatively raise a new problem in the field of vestibular function. The following considerations may help to form an estimate of the accuracy with which the vestibular data must represent the actual head movement in order to make possible the localizations which were obtained in this experiment.

As can be seen from the figures given above, in this experiment the loudspeakers close to the original one are 4½° apart. Roughly this is also true of the others. Since the contact points on the switch are 3° apart, the shift of the sound direction actually presented is 1½ times as great as the corresponding displacement of the head. In the experiment first reported the shift of the sound is twice as great as the displacement of the head, and the image of the sound appears hori-

[7] With respect to the achievement of sound localization these two possibilities are equivalent.

zontally opposite the loudspeakers. Another case which we may consider here is the one in which the image of the sound is directly overhead. As will be seen below, in this case the shift of the sound actually presented is as great as the displacement of the head. In short, if the ratio of the displacement of the sound to that of the head is 1, the sound image is straight above; if the ratio is 1½, the image is elevated by 60°; and if the ratio is 2, it is horizontal. A simple calculation shows that in the present case, where the ratio is 1½, an overrating of the actual displacement of the head by 50 percent due to insufficient sensory data would make this ratio 1, and accordingly the sound image would appear straight above; whereas an underrating of the actual displacement by 25 percent would make the ratio 2, and the sound would appear horizontal. Thus, in this case an underrating of a displacement of the head affects the perceived direction more strongly than an overrating of the same order.

An examination of the table shows that practically all the inadequate localizations which occurred were too low. This may be explained by the asymmetry with which inaccurate sensory data for the displacement of the head affect the perceived direction. In spatial terms, an inadequately determined direction may thus have a better chance of being heard below the desired direction than above. However, the simple assumption that with vision excluded the displacement of the head is underrated rather than overrated could also account for this finding, and in the experiment to be reported next, this latter explanation seems to be the proper one.

A sound image vertically above the observer was synthetically produced. In this case the arrangement of the loudspeakers is especially simple. A sound direction vertically above is characterized by the fact that the sound remains within the median plane when the head turns about a vertical axis; the lateral angle remains 90°. Thus, during such a displacement, the lateral angle does not change at all, and the rate of change is, of course, also zero. In order to duplicate these conditions in a synthetic experiment, the sound actu-

ally presented in the horizontal has to shift in such a way as to maintain a constant lateral angle of 90°. This is the case, when the actually presented sound is always either straight in front or straight in back of the observer's head. One achieves this by giving the loudspeakers angular distances of 3° from each other, the same as the angular distances of the contact points on the switch. While the observer's head is being attached, the switch is in a position which connects the loudspeaker straight in front or back. Under these conditions the actually presented sound is located in the median plane for any position of the head. When the head is thus attached, the experiment can proceed in the same way as the one reported above.

Of the 10 subjects who took part in the experiment, two were unable to localize the sound in the region overhead, even under favorable conditions. One of them always heard it in the rear about 60° high, and the other one horizontally in front or in back. The results for the other 8 subjects are as follows: 90°, 90°, 90°, 80°, 75°, 75–70°, 70°, 70°; here 90° represents the vertical, the other angles refer to directions slightly in back of the vertical. Thus, when blindfolded, 3 of the observers heard the sound in the desired direction, and the others localized it from 10 to 20° too far back. When the blindfold was removed, they all had the image of the sound vertically above. Again all deviations from the desired direction corresponded to an underrating of the displacement of the head; and since here misrepresentation of the actual displacement of the head in either sense leads to the same amount of deviation in spatial terms, we can safely adopt underrating of the displacement as the explanation. A simple calculation shows that an underrating of the displacement by 25 percent accounts for a deviation of 20° toward the back, which is the maximum which occurred in this experiment. In another respect this experiment confirms the result of the preceding one: There are subjects who can adequately localize merely on the basis of vestibular cues derived from acceleration, and with the others inaccuracies are restricted to what corresponds to a misrepresentation of the

displacement by 25 percent. With the aid of vision, moreover, passive displacement is just as effective as active head movements.

The distinctly directed deviations from the desired direction which were found in these experiments seem to me of particular interest for the theory of sound localization. They obviously cannot be regarded as the result merely of an indistinct perception. In a purely geometrical sense there are two ways in which the elevation of a direction can be characterized by the behavior of the lateral angle during a displacement of the aural axis. One way is that the amount of change of the lateral angle referred to the displacement of the head as such characterizes the perceived direction, as has just been discussed. The value of this quotient would here directly determine the process which corresponds to the perceived direction.

The other way was brought out in the discussion of the maximal differences in the change of lateral angle which may occur with different elevations of a source of sound: the range within which the lateral angle of an elevated direction can change depends upon the elevation above the plane in which the aural axis is displaced. The greater the elevation of the sound direction above this plane, the smaller the range of variation of the lateral angle. For a head movement about a vertical axis, for instance, the lateral angle for a direction elevated 60° above the horizontal plane varies only within a range of 30°. An extreme case is the vertical direction; its lateral angle does not vary at all. No matter how far the head is displaced, the smallest value which the lateral angle of an elevated direction can assume is obviously the same as its angle of elevation above the plane in which the aural axis shifts. Thus the limiting value of the lateral angle can directly define the elevation of a direction. This constitutes the other way in which the elevation of a direction might be characterized by the behavior of the lateral angle.

It is clear from the outset that the second principle cannot alone be responsible for sound localization. Long before head movements produce the limiting value of the lateral angle, accurate sound localization may be achieved. The question is rather whether the second principle—which *geometrically* is strictly connected with the first—plays a secondary rôle in sound localization. In other words, are geometrical relations so fully represented in the process of sound localization that, because the first principle holds, the second—its geometrical corollary—must also hold?

In this connection, the most interesting case is that of the vertical direction because here the second principle could yield as strict a determination as the first. If we say that a sound is heard above because its lateral angle does not change at all but remains 90° throughout, this statement as such could be interpreted in terms of both principles. In terms of the first we might say that the lateral angle does not change, and in terms of the second that it remains 90°. Either one of the arguments would suffice to characterize the direction. But the question is whether both principles are actually involved in the process. Here the experiment which we have just reported is decisive. When visual cues for the displacement of the head are excluded, the perceived direction is often definitely too low. Since these deviations always disappear when the blindfold is removed, they seem to be due to the insufficiency of available data for the displacement of the head, more specifically to the fact that the displacement of the head is underrated. The sound is heard at an elevation of 70° or 80°; and yet the actually presented sound is given in the median plane, *i.e.*, with a lateral angle of 90°. If the second, the limiting, principle were effective, this could not occur.

We must therefore conclude that this second principle plays no rôle in the process of sound localization. It is merely the *amount* of the change of the lateral angle referred to the displacement of the head which is effective in determining the elevation.

## III. The Selective Principle

Before further experiments are reported in which sound localization is based on *visual* cues, it seems advisable to discuss briefly a general selective principle which is implied in all sound localization. This principle becomes particularly obvious in experiments in which sound directions are synthetically produced. Why is it, one may ask, that the desired direction is heard instead of the actually presented sequence of directions? Obviously because this direction is at rest whereas the actually presented sound shifts in space while the head is being displaced. Apart from this selective principle, the actually presented sequence of directions and the direction which is perceived are, according to our previous discussion, entirely equivalent. They produce, with the head movement in question, exactly the same temporal pattern of binaural stimulation. We know that almost every lateral angle can be represented by a number of different directions, and that a given sequence of lateral angles can thus be represented by a nearly endless variety of patterns of subsequent directions. But in a given case this sequence of lateral angles, no matter how it is produced, gives rise to one percept only, that of a *stationary* direction which is compatible with the sequence.

For the sake of simplicity we shall confine this discussion to sequences in which each succeeding direction is contiguous with its predecessor, *i.e.*, cases where the presented sound shifts steadily and, moreover, in a plane. Even in this simplest case it may be hard to visualize the spatial relations which result from the circumstance that both the presented sound and the head are displaced. The task becomes easier if one chooses the head as frame of reference. With reference to the head, all directions which realize a given sequence of lateral angles are displaced during the head movement, the actually perceived one as well as the others. Objectively, the percieved direction is distinguished from the others only

by one fact. The perceived direction is covariant with further objects which are given in our environment by other senses and which are, of course, also displaced with reference to the head when the head is being turned. We thus arrive at the following formulation of the principle of rest which I believe to be the most adequate: Of all the directions which realize the given sequence of lateral angles, that one is perceived which is covariant with the general content of surrounding space.

It is significant that the principle of rest must apparently be regarded as the limiting case of a more extensive principle which, although not overt in ordinary sound localization, seems to indicate the general way in which the selection is made. This broader principle was demonstrated in synthetic production of sound directions when sequences of lateral angles were presented for which no stationary direction existed. It was named the principle of least displacement. If a sequence of lateral angles is presented to which no stationary direction corresponds, the sound is perceived in the region where it has to undergo the smallest displacement in space while realizing the given sequence of lateral angles.

Suppose a sound direction vertically above is to be synthetically produced by means of a sound that shifts in the horizontal plane. In this case the arrangement of the head switch and the loudspeakers must be such as to leave the actually presented sound in the median plane of the head. This is the condition which must be fulfilled if the desired vertical direction is to be obtained. Consequently the angular distances of the contact points and of the loudspeakers will have to be the same, and if the head is fixed to the switch in such a way as to connect at the start the loudspeaker straight in front, the sound travels about during the head movement so that it is at any moment straight in front of the head. If, on the other hand, the head and the switch are connected, when the loudspeaker in question lies, for instance, 25° to the right, the sound will, during the head movement, always remain 25° to the right of the median plane of the head, *i.e.*, at a lateral angle of 65°. No stationary direction cor-

responds to this sequence. The only case in which a sound at rest can be produced with a constant lateral angle is that in which this angle amounts to 90°. Here, it will be remembered, the sound image appears in the direction of the axis of the revolution of the head. When the lateral angle remains constant at 65° the sound is perceived at an elevation of about 65° to the right of the median plane, that is as close as possible to the axis of the displacement. Here it moves about in correspondence with the displacement of the head.

It is obvious that at any lower elevation the shift in space necessary for the realization of the given sequence of lateral angles would be greater. The closer the sound to the axis of the revolution, the smaller is the displacement of the sound image in space. In accordance with the principle of least displacement, the sound is actually perceived in maximal proximity to that axis.[8]

## IV. Rotation of the Visual Field

In Section II of this paper, experiments were reported in which both proprioceptive stimulation resulting from active head movements and visual cues for the displacement of the head were excluded. Quite substantial achievements in sound localization were shown to occur when only vestibular stimulation was thus admitted. If visual cues for the displacement of the head were added to the vestibular stimulation, they improved the achievements and even made them optimal The question will now be examined of whether visual data *alone* can represent the displacement of the head in sound localization.

In order to eliminate vestibular stimulation, it seems to be necessary to keep the observer physically at rest and bring about a psychological state of motion by means of induced ego-movement. When an observer is placed inside a revolving screen he will, after a while, no longer perceive the surrounding screen in motion; rather he will feel himself rotating in a direction opposite to the objective movement of the

[8] Another experiment in which the principle of least displacement determines the result is reported in 1 §22.

screen. The question is, of course, whether with respect to sound localization the optically produced state of phenomenal movement is functionally equivalent to real movement; in other words, whether optical stimulation can replace vestibular stimulation.

The revolving screen used in this experiment was made of white cloth which hung down from the edge of a large wheel fixed in a horizontal position underneath the ceiling. It thus formed a hollow cylinder 43″ in diameter. On the inside it showed vertical black stripes on a white ground, 2¼ inches wide and 6¾ inches apart. Wheel and screen were turned by a motor at various speeds.[9] When an observer was placed in this small compartment he soon felt himself turning in the opposite direction, and soon afterwards the screen appeared to be at rest. When now the observer looked down at the floor, he saw it turning in his own direction and at the same rate. The spontaneous transition from rest to motion was often quite disagreeable, but once the screen had come to phenomenal rest, as a rule no giddiness was felt. When the state of complete ego-movement was attained, there was never a spontaneous change back to the first state in which the observer is at rest and the screen moves. The transition to ego-movement was facilitated when some object was placed inside the screen upon which the observer could let his eyes rest. The observer was instructed to look steadily at a vertical rod which was placed before him and as close as possible to the screen. Soon after the screen began to turn the rod showed induced movement in the direction opposite to the movement of the screen; and gradually the observer felt himself joining in this motion until after a while, with the cessation of the phenomenal movement of the screen, the state of induced ego-movement became complete. Thus the momentary giddiness which often accompanies the transition to ego-movement could be eliminated.

A sound which remains in the median plane when the head is turned will be perceived in the direction of the axis of the

[9] One revolution in 7 seconds proved to be a comfortable rate for the screen which was used.

displacement. If this axis is vertical, a sound objectively straight above the head would fulfill this condition. But the same condition is satisfied when a sound which shifts about in the horizontal plane remains throughout in a median position with reference to the head. If induced ego-movement can be substituted for an actual displacement of the head, the same effect can be obtained in an experiment with the revolving screen. A loudspeaker is placed at some distance beyond the screen straight in front of the observer whose head is kept in a constant position by a chin-rest. When now during induced movement a sound is presented in the loudspeaker, the above conditions for hearing a sound vertically above are given: the head of the observer 'turns' about a vertical axis, and at the same time the sound remains always in a median position with reference to the head. In this stiuation the sound is actually heard vertically above. Thus it is shown that induced ego-movement can be substituted for a physical displacement of the head.

The great simplicity of this form of the synthetic experiment lies in the fact that instead of two displacements, that of the head and that of the sound, only the screen is actually moving. In the state of induced ego-movement the revolving screen represents the resting space, and all the objects which are physically at rest are therefore represented in the same state of motion as that in which the observer who is also physically stationary feels himself. The floor of the room and the rod in front of the observer are, as reported above, perceived in this state of motion. The loudspeaker would move about in the same manner, if the observer could see it. Had the distribution of motion and rest which the observer perceives in the state of induced ego-movement been objectively given, this experiment would be an exact duplication of an ordinary synthetic experiment: the head of the observer is displaced about a vertical axis, and the source of sound undergoes a corresponding displacement, so that it always remains in a median position.

Of 15 observers who participated in this experiment, 3 were unable, even under ordinary circumstances, to localize

a sound presented objectively overhead. The remaining 12 perceived the sound as vertically above the head, either immediately or shortly after the beginning of the presentation. In the few moments during which it stayed in the horizontal plane it seemed to move about with the observer. When a sound was presented exactly to the side, it likewise travelled phenomenally with the observer, remaining always in the same position with reference to him. Yet, in this case, it remained in the horizontal plane. A sound direction which coincides with the aural axis is strictly determined by the binaural cues alone, just as the pole of the globe is sufficiently defined by its latitude, and needs no longitudinal determination. No other direction produces the same binaural cues, and thus no other direction could replace it.

If the loudspeaker was placed obliquely to the side, say, 30° from the median position, the sound was heard obliquely above where it seemed to move about with the observer. From the above analysis, it is clear that this is an analogue of the experiment on the principle of least displacement which was reported in Section III. All these observations can be made in immediate succession, when the loudspeaker is slowly carried around from a position straight in front of the observer to one on his side. The sound which is at first heard directly overhead descends slowly to a horizontal position.

There is still another experiment which can be performed with induced ego-movement. It corresponds to the first experiment reported in this paper, where a sound actually presented in front was perceived in back, and vice versa. This reversal of front and back was brought about when the actually presented sound direction was displaced by twice the angle of the displacement of the head. When induced ego-movement is substituted for a physical displacement of the head, the same relation of the displacement of the head and of the presented sound can be brought about. During induced ego-movement, it will be remembered, the observer and a stationary sound will be felt in the same state of motion. In order to give the sound *twice* the displacement of the observer it must in addition be physically displaced in the direction in

which the observer feels himself turning, and this at a rate which is equivalent to the rate of the observer's subjective motion. If we assume that the subjective rate of motion is (in the opposite direction) the same as the objective rate of the screen, the sound must be objectively displaced with an angular velocity equal to the angular velocity of the screen, but in the opposite direction. Thus, in order to obtain by means of induced ego-movement, the reversal of front and back, we must rotate the screen as well as the source of sound about the observer, the screen in one direction and the sound in the other direction, but both at the same rate.

In order to rotate the sound about the observer, a small dynamic loudspeaker with permanent magnet was fixed to the end of a long arm which could be swung in a circle about the observer. The seat for the observer was fixed on a vertical column consisting of 1½″ steel tubing mounted on a heavy iron base. The arm which carried the loudspeaker was rigidly fastened to a wooden block which turned on two ball-bearings about the column between the base and the seat. The block carried a pulley by means of which the arm was turned around. The outer end of the arm was bent up to bring the loudspeaker to the level of the observer's head. Mounted on the base was a reduction gear motor which, through a rubber belt, set pulley, block, arm and loudspeaker in rotation. A heavy lead fly wheel on the main shaft of the motor kept it running at a constant speed. This whole device was placed within the revolving screen with the arm projecting out under the lower edge. The distance from the head of the observer to the loudspeaker was 150 cm. The current for the loudspeaker was carried to the moving parts by two iron wires which, fixed to the block, dipped into circular grooves filled with mercury which were cut into a bakelite block fastened to the base.[10]

The procedure in an experiment was the following. After the observer had been placed on the seat inside the screen, the loudspeaker was made to turn around him. When the proper rate of rotation was reached, the screen was set in motion at a rate previously determined to match the speed of the revolving loudspeaker. The screen turned to the right while the loudspeaker shifted to the left. Soon the observer would find himself in rotation to the left, while the screen appeared to be at rest. Then the sound was presented. Had the observer localized it in the loudspeaker from where it actually came, he would have perceived it moving around in the same direction in which he felt himself moving, only twice as fast. But this never happened. Rather the localization

[10] The writer is indebted to Prof. E. B. Newman for designing this apparatus.

of the sound was always reversed with respect to front and back. The observer heard it straight in front when the loudspeaker actually passed the median plane in back of him, and vice versa. For some observers the sound was definitely at rest, and they felt themselves passing it as they turned about, just as if they were turning on a revolving chair and were passing again and again a stationary source of sound. Others could not decide whether the sound moved or remained stationary. Often it was heard moving slowly to the right, that is in the direction opposite to the subjective rotation of the observer and the objective displacement of the loudspeaker. Irrespective of such differences, all 20 observers who took part in this experiment showed the reversal of front and back. When asked "How does the sound behave with respect to you?" all observers reported that it shifted to the right, that is, in the direction opposite to the actual displacement of the sound. This fact is necessarily connected with the reversal of front and back. The binaural cues strictly determine the position of the sound with respect to left and right. Thus the sound is invariably heard on the left when the loudspeaker is on the left and on the right when the loudspeaker is on the right. Suppose now that the loudspeaker travels objectively, for instance, from a position on the left side to a position in back. Since at the same time the phenomenal sound moves from the left to the front, the reversal of the direction of revolution follows.

It may be well to consider how the general selective principle which was discussed in Section III applies to this experiment. Are the directions in which the sound is perceived covariant with the general content of surrounding space? More specifically: Is the sound in its reversed positions covariant with the screen? Let us call the physical position of the loudspeaker when it is straight in back of the observer position 1. In this position the sound is heard straight in front. When in its objective revolution to the left the loudspeaker has shifted by 90° from its position in back to the right side of the observer (position 2), it is actually heard in this position. Meanwhile the screen, too, has been

displaced by 90°, and since its rotation is to the right, the part, which was in front when the loudspeaker was in position 1, is now also to the right of the observer. During the next quarter revolution the loudspeaker shifts from the right side to the position straight in front (position 3). At the same time the perceived direction changes from the right side to the position straight in back. Yet the part of the screen which just lay in the direction 90° to the right is now likewise straight in back, because the screen, too, has undergone another displacement of 90°. In the same fashion positions of the screen and of the perceived sound continue to coincide for the remainder of the revolution. Thus we see that the direction of the reversed sound remains always the same with reference to the screen. The selective principle applies to this experiment.

In the state of induced ego-movement the screen appears at rest and consequently the perceived direction should also be perceived at rest. This was actually the case with those observers who reported that they seemed to pass a stationary sound in their apparent rotation. Others, however, heard the sound moving in the direction opposite to their own rotation. This observation cannot be explained in terms of our previous discussion. Purely geometrical considerations lead to the conclusion that the reversed sound direction is covariant with the screen. If the sound is heard in a reversed direction, as is actually the case, this direction must always coincide with a particular part of the screen. However, if in spite of this covariance a moving sound is perceived instead of a stationary one, we have a new problem before us. This problem belongs to the psychology of movement rather than to sound localization.

A last point remains to be mentioned. The reversal of front and back can often be shown to occur when the observer is not yet in the state of induced motion. The stimulus condition alone which eventually results in phenomenal movement of the self apparently suffices to produce the reversal of front and back. Here it should be remembered that, as far as visual stimulation is concerned, the situation in which a

subject is actually turning in a stationary screen is entirely equivalent to the situation in the present experiment where the screen turns and the observer remains physically at rest. It seems that a visual stimulation, which is equivalent to that resulting from rotation of the subject, produces the corresponding effect on sound localization, and that this effect occurs regardless of whether the phenomenon of ego-movement accompanies this stimulation.

In Section III two alternative formulations of a general selective principle were offered. The evidence which has just been cited favors one of them. It now seems doubtful that phenomenal rest as such is the factor which distinguishes one direction from all the others which are compatible with the presented sequence of binaural cues. Covariance with the general content of surrounding space, however—whether it is perceived at rest or not—can still be regarded as decisive for the selection of the perceived direction.

## Summary

Experiments of synthetic production of sound directions have been reported which show that either vestibular cues or visual cues can replace head movements as such. In one group of experiments the blindfolded subject localized the sound while he was passively turned on a revolving chair, and in the other group the subject observed the direction of sound while seated inside a revolving screen. The results indicate that (*a*) fairly accurate representation of the actual displacement of the head is furnished by vestibular stimulation and that (*b*) visual stimulation, equivalent to that which actual displacement of the head would give, suffices to determine the direction of sound.

(Manuscript received May 14, 1940)

# STUDIES OF ABNORMAL BEHAVIOR IN THE RAT. IV. ABORTIVE BEHAVIOR AND ITS RELATION TO THE NEUROTIC ATTACK[1]

BY NORMAN R. F. MAIER

*University of Michigan*

## Introduction

When a blast of air was used to force rats to jump to a card that they had previously been trained to avoid Maier (1) found that they frequently exhibited a convulsive attack which was described as the 'neurotic pattern.' Control tests showed that the air alone would not release this behavior in most of the rats reported in this study and it was therefore believed that the conflict between excitation (produced by the air) and inhibition (produced by the negative card) was a basic factor in the appearance of the attacks. A later study by Maier and Glaser (2), showed that exposure to various forms of auditory stimuli (jingling of keys, sound of rushing air, electric bell and electric buzzer) produced attacks in a large percentage of animals in the colony but the addition of barriers which restricted the movements of the animals and the addition of a conflict based upon jumping to the negative card greatly increased the occurrence of attacks in a group of rats. Maier and Glaser, therefore, concluded that even exposure to direct stimulation could be regarded as a conflict situation since the auditory stimulus in such cases gave rise to no specific response but was excitatory in nature and the enclosed space restricted general behavior and served as the inhibitory aspect of the situation. This interpretation was supported by the fact that attacks increased as the conflict between excitation and inhibition became more and more specific. It was also found that there was little correlation between the frequency of attacks in the

[1] This study is part of a program of work supported by a grant from The John and Mary R. Markle Foundation of New York City. The author wishes to acknowledge his indebtedness and to express his appreciation for this aid.

various test situations even when the auditory aspect of the situation was constant. This lack of relationship suggests that the cause of the convulsion is perceptual rather than sensory in nature, and rules against any theory which tends to reduce the problem to a stimulus-response level.

If we consider the basic mechanism underlying the abnormal behavior to be one of conflict one may expect the abnormality to disappear if a mode of behavior which resolves the conflict can be established. Maier (1) pointed out that many rats develop peculiar modes of jumping, which he called 'abortive jumping,' when they were required to choose the negative card and suggested that this manner of responding to the situation reduced the tendency to have an attack. It is the purpose of this investigation to study further the rôle of abortive behavior and see whether its development and its prevention might not be related to the disappearance and reappearance of the neurotic pattern. If abortive behavior can be set up as an alternative to the neurotic pattern its development will in effect constitute a psychological cure for the convulsive disorder. Previous observations have shown that abortive behavior will appear in the Lashley jumping apparatus [2] whenever the probability of punishment for a choice becomes fairly great and the animal is forced to react. Thus if neither card is consistently latched in a two-window discrimination test (the no-solution problem) or if only one window is present and contains the negative member of a training pair (the one-window situation) the rat is likely to stereotype an abortive jumping response, providing it is forced to respond. Both electric shock and blasts of air have been used to break resistance to jumping in such frustrating situations. [See Maier (1), and Maier, Glaser, and Klee (3).]

[2] The Lashley jumping apparatus consists of a jumping platform from which the rat is required to leap at and strike a card covering a window in a large vertical screen about ten inches from the jumping platform. By latching the card the rat may be punished by a bump on the nose and a fall into a protecting net below. When the card is unlatched the rat may knock over the card and be rewarded by gaining access to food placed behind the window. If the animal is given a choice between two cards and is rewarded for leaping to one and punished for leaping to the other, a preference for one of the cards is readily established.

The abortive jumping responses thus far observed are as follows:

1. Instead of striking the card with the nose the rat strikes the card with the side of the body and thus goes to either the right or the left of the card. In a more exaggerated form the rat leaps to the right or left of the card and strikes the screen lightly with its four legs or with the side of the body and then falls to the net. In some cases the rat may leap twelve inches to the side of the card. We will refer to these responses as right or left abortive jumps.

2. If no screen is present above the cards to which the animal must jump, a leap above the card is fairly common. In some cases the animal leaps over a foot above the card. (See Fig. 1.) A rat may also combine this form of response with the former and jump above and to the side of the card. In the pure form this response is referred to as 'abortive over' and in combination as 'over right' or 'over left.'

3. Since the stimulus card is placed behind the screen some rats become proficient in leaping lightly and catching hold of the window frame with the forepaws where they hang until rescued or until they fall. This form of abortive jump is fairly difficult to discourage and is referred to as 'hanging on.'

4. In a few cases the rat develops a light jump which carries the animal to the card but which is inadequate for gaining the ledge even when the card falls over. In some animals this response is so slight that the animal does not reach the card at all, so that the response becomes, in effect, a jump directly to the net. This form of response is the most difficult to recognize since one cannot always detect the light jump when the rat reaches the card. It is also the most difficult form of adjustment to prevent. We speak of this form of abortive response as the 'light jump.'

5. The construction of the starting box was such that a rat could go to the front of the box and turn its body and climb to the top. This abortive response will be referred to as 'escape behavior.'

It is characteristic that a specific form of abortive behavior becomes stereotyped. Each animal has his individual form and continues it unless interference is introduced. Motion pictures illustrating abortive jumping have been made by Maier (4).

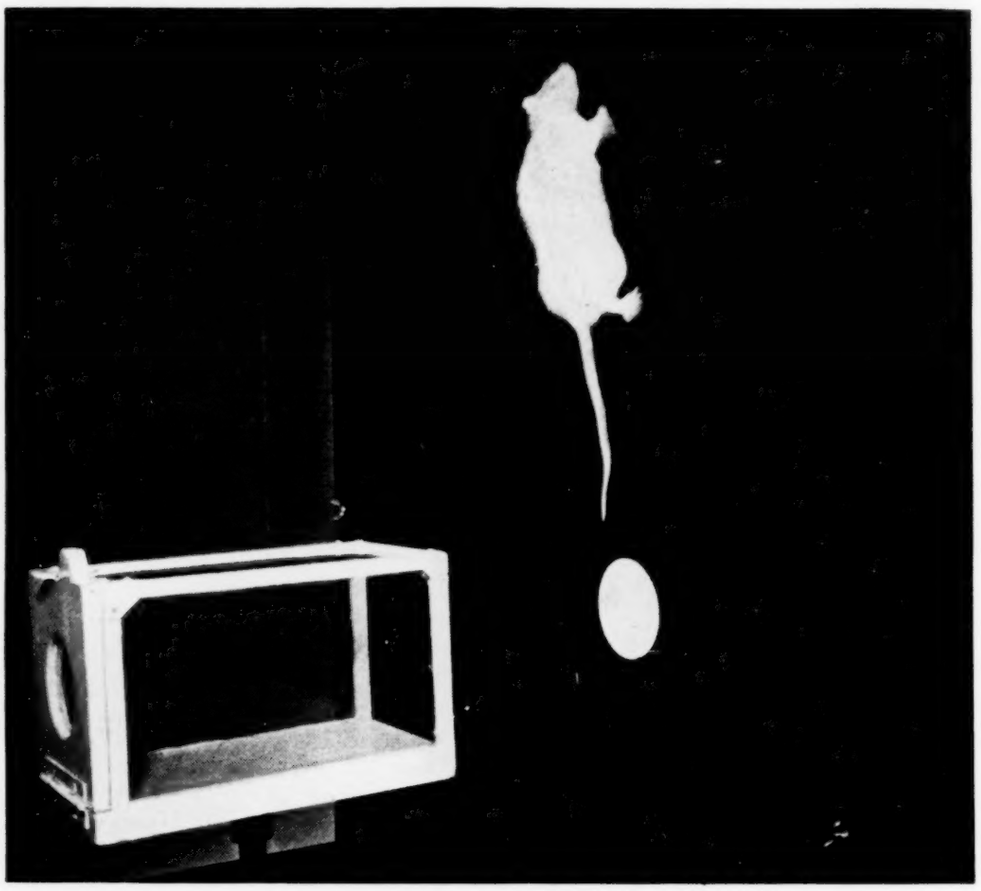

FIG. 1. The 'new' jumping apparatus with an illustration of abortive jumping. The enclosed box is the jumping stand from which the rat is required to leap at the card placed in the window of the vertical screen. This screen is removeable so that a two window screen may be substituted for the one shown in the photograph. Thus the same apparatus may be used for the one- and the two-window problems. A removeable cardboard screen may be fastened above the windows and so prevent the animal from leaping over the card. In the illustration this screen is not in place. The photograph was taken while a rat was executing an abortive jump well over the negative stimulus card. The air used to drive the rats is led to the jumpingbox by a rubber tube shown to the left in the picture and just below the door of the box. The air is released from a metal tube (slightly flatened at the end) which is bent at right angles so that the air strikes the animal from behind.

## METHOD AND PROCEDURE

In order to study the rôle of abortive behavior, rats which had previously been extensively studied were used. Each

animal therefore had already learned to discriminate between the negative and positive cards and each had a record of attacks in the one-window situation. Because each rat had its individual characteristics, conditions in this experiment were manipulated differently in each case. For this reason it is necessary to present the record of each animal separately.

Before one can conclude that abortive behavior is the cause of the disappearance of the neurotic pattern one must exclude other possible causes, such as differences in the amount of exposure to air stimulation, changes in the animal's physiological condition, and adaptation to the situation. It is therefore desirable that the neurotic pattern not only be caused to disappear but be made to reappear when abortive behavior is abolished. This is not always an easy matter since a highly susceptible animal may have days on which attacks cannot be produced and fairly stable animals may have attacks when conditions have not been altered. Because fluctuations in susceptibility are common and because fairly permanent changes in the animal may occur for reasons unknown to the author it was necessary to continue these experiments over a considerable period of time.

Three types of situations were used in these experiments. (1) *The discrimination test* which consisted of a presentation of the positive and negative cards of the original training series (black circle in white background and a white circle in a black background, respectively) and then requiring the animal to make a choice.

(2) *The one-window situation* in which only one of the above cards was presented and the rat was required to jump to the card irrespective of whether it was the negative card (and therefore latched) or the positive card. Since the screen containing the windows was removable the discrimination situation (containing two windows) could be changed into the one-window situation by merely exchanging screens.

(3) *The two negative card test* which was like the discrimination test except that the two cards presented were alike and negative. Since abortive jumping is highly specific a given abortive response would not have the same relation to the

two negative cards as to the one negative card presented in the one-window situation.

In all of these test situations air was turned on as soon as the rat was placed in the jumping box. The air was introduced behind the rat and was constant in intensity.

The experiments were performed on two different apparatuses. The *old* apparatus was used for all tests preceding the present experiment. Most of the data in this study were obtained with the *new* apparatus, but frequently the old apparatus was used to test possible differences in effect. The new apparatus was like the old in all respects except that (1) the starting box was so constructed as to permit the 'escape' form of abortive behavior and (2) the screen above the cards was removable and when detached permitted abortive jumping over the stimulus cards. The new apparatus thus made abortive behavior more accessible. Figure 1 shows the new apparatus with the screen above the cards removed and with the one-window test in progress.

## Results

### *Rat 32M*

*Previous History*

This rat was trained to discriminate between a black circle on a white background and a white circle on a black background. When the discrimination was complete a series of new pairs of cards (in which the size and brightness of the figure and the brightness of the background were altered) was presented and the animal was forced (by means of air) to choose between the members of each pair. In some cases the previous training transferred, in others the training was inadequate to cause one of the members to be preferred. The first attack occurred in this rat when confronted with a critical pair in which no preference for one of the members was expressed (a non-equivalent pair). After further experimentation attacks also occurred when a preference for one member of a critical pair was shown (equivalent pairs) and soon even the original training pair resulted in attacks. (For further details see Maier (1).)

After a two-months interval the rat was transferred to the 'one-window' situation and attacks occurred quite regularly. On 11 different days it was presented with the one-window situation and attacks occurred on 8 occasions (73 percent). This rat thus showed attacks regularly over a considerable period of time.

*Cure*

After another month of rest a detailed study of this rat was begun to see if it could be cured. It was transferred to the new apparatus and because the screen above the cards was eliminated it was now possible for the rat to jump abortively by leaping above the stimulus card.

The training transferred to the new apparatus and the animal made a perfect score on the discrimination problem. The one-window situation was then introduced and almost immediately abortive jumps to the right or above the card appeared, particularly in response to the negative card. After the first day the animal was also tested on the old apparatus and on these trials it abortively jumped to the right of the card. (The screen prevented 'abortive over' patterns.) Thus the abortive manner of jumping transferred to the old apparatus. For 4 days the rat was given from 5 to 10 daily trials on each apparatus and clear-cut abortive jumping occurred on over 85 percent of the presentations of the negative card and on 40 percent of the presentations of the positive card. On other occasions it tended to jump lightly at the cards.

Never before had this rat been in the situation this much without having an attack and it was believed that the abortive manner of jumping was responsible for the difference.

To disturb the abortive pattern of jumping the rat was now returned to the discrimination problem on the old apparatus and on 5 trials it chose the positive card without evidencing abortive jumping. Then two negative cards were introduced. On viewing the card the rat became very nervous and after 15 seconds' exposure to air it jumped from the jumping box and had a violent attack.

During this and the next 14 test periods scattered over 32 days the rat was exposed to the discrimination problem on 7 occasions, the one-window situation on 8 occasions, and 2 negative cards on 12 occasions on either one or both of the apparatuses. The number of trials on each situation was from 5 to 10 and several situations frequently were presented on the same day. A single test occasion or exposure to a situation thus involves at least 5 trials.

Table 1 shows the number of attacks in each situation.

TABLE 1

DISTRIBUTION OF ATTACKS DURING 5TH THROUGH 19TH TEST PERIODS FOR RAT 32 M
(Numerator shows number of attacks, denominator shows number of tests.)

| Situation | Apparatus | | Total |
|---|---|---|---|
| | Old | New | |
| Discrimination | 1/3 | 1/5 | 2/8 |
| 2 negative cards | 6/8 | 0/4 | 6/12 |
| One-window | 0/0 | 0/7 | 0/7 |
| Total | 7/11 | 1/16 | 8/27 |

The data obtained with the new and old apparatuses are presented in separate columns and the results are expressed as the ratio between the number of attacks and the number of exposures in each situation.

It will be seen that the one-window situation which had previously been effective in producing attacks now failed entirely (no attacks on 7 occasions). In this situation abortive jumping became very stereotyped during this period as the animal jumped well over the negative card. On both apparatuses it jumped through the positive card quite consistently during these trials, thus showing good discrimination.

The presentation of two negative cards, however, resulted in 6 attacks on a total of 12 exposures. All of these attacks occurred on the 8 tests given on the old apparatus, while 4 such tests on the new apparatus were ineffective. Instead the rat jumped abortively over the right member of the pair of negative cards. Of importance also is the fact that the attacks obtained always occurred on the first presentation of

the 2 negative cards. In the case of three of the attacks, exposures to the one or both of the other situations preceded and in the case of two of the attacks the two-negative-card test was given at the outset. This distribution of attacks shows that (1) the method of piling up of strains and (2) the accidental association of a certain test with days on which the animal was susceptible are not responsible for the attacks, but rather the attacks are associated with the situation in which abortive jumping had not been developed.

The two attacks which occurred in the discrimination problem were very mild in that the animal showed an initial burst of activity and then immediately came to rest with no apparent stupor thereafter. These attacks occurred respectively on the 3rd and 7th trials of the 7th and 8th day of testing and just at the time when abortive jumping was becoming associated with the discrimination problem. The appearance of abortive jumping in this situation indicates a breaking up of the well-established discrimination habit.

In all of these situations air was applied to force the animal to react. The attacks all occurred after 10 to 25 seconds' exposure to air. In each situation in which no attacks occurred the duration of air exposure was definitely greater than 25 seconds (and of the same intensity). This was true of at least some of the trials on every test day. Therefore it cannot be contended that attacks occur because of increased exposure to air on crucial test days.

The foregoing evidence indicates that as far as the one-window situation is concerned the animal is cured. Although the one-window test was not given on the old apparatus during the period covered by Table 1, it should be recalled that for 3 days preceding this period the rat showed a right abortive response in the old apparatus and it was not until this was disturbed that attacks occurred. That this cure is not due to a loss of the pattern occurring for reasons unknown to us (such as a changed physiological condition) is shown by the fact that attacks can still be elicited by other means. That the abortive behavior is responsible for the adjustment is supported by the finding that the cure applies

only to the situation in which the abortive behavior appears as a regular and highly specific mode of behavior.

To further substantiate this contention we should be able to eliminate the abortive behavior and so produce attacks in the one-window situation on the new apparatus.

*Reinstatement of Attacks*

We therefore placed a screen above the cards on the new apparatus, thus barring the stereotyped abortive response over the cards, and presented the rat with the one-window situation. On the first trial with the negative card in place the rat showed a violent attack. The attack followed 15 seconds' exposure to air. On all preceding days in which attacks failed to occur exposure to air surpassed this amount, the average maximum exposure per day being 31 seconds. This shows that the one-window situation is still effective if the stereotyped abortive behavior is successfully eliminated.

During the following 12 days, 9 test periods were given. On 8 of these test periods the one-window situation was presented; on 3 of these periods the screen was present, on 2, no screen was present, and on 3, tests were made both with and without the screen. However, no attack occurred in any of these tests, but instead the rat altered its abortive behavior by jumping to the right of the card as it had done on the old apparatus. When such abortive behavior was blocked, it jumped to the ledge and hung there with the front paws. Similar behavior was shown when the two negative cards were presented on two of the 9 test days.

On each of the 9 days there were trials in which 35 to 80 seconds of air (average 54 seconds) were required to produce a jumping response. Because all neurotic attacks have occurred in this rat with less than 30 seconds' exposure to air it cannot be claimed that exposure to air during this period was inadequate.

Because the rat now freely exhibited various kinds of abortive behavior it was believed that retraining might eliminate these adjustments. The screen was therefore removed and for two days the original discrimination was given. On

each trial the rat jumped normally to the positive card when it appeared on the right side but jumped over the positive card when it appeared on the left side. Resistance to jumping was marked and air had to be applied from 1 to 2 minutes. Thus the rat in addition to its discrimination habit now had a position habit in the sense that it jumped differently to the right and left sides.

On the following day the discrimination problem was again presented. On the first trial with the positive card on the right and after 30 seconds' exposure to air it had a mild attack. The attack thus occurred on the side where no abortive behavior appeared.

The next day abortive behavior disappeared and the rat jumped normally in the discrimination problem. Following these normal responses to the discrimination problem it was presented with two negative cards. After 22 seconds' exposure to air another mild attack occurred.

No attacks occurred on the next day even with the screen in place. Rather the rat jumped to the right of the cards when two negative cards were presented as well as when the discrimination problem was presented. Thus abortive behavior was again present.

We now decided to stereotype the abortive behavior to see if other forms of abortive behavior would then disappear. During the next 5 test periods the rat was given 5 trials in the two-negative-card test with no screen above the cards. At first the rat jumped abortively by jumping to the right of the cards, or over the right card, or both over and to the right, but gradually the jump over the card dominated until on the fifth day the rat consistently jumped above the card (often by as much as 8 inches). Each day trials occurred in which the rat was exposed to air for as long as 65 to 130 seconds (average 102 seconds) before it jumped.

With the response of jumping over the card stereotyped, the screen above the card was introduced the next day. On the first trial, with two negative cards in place, a characteristic attack occurred after 29 seconds' exposure to air.

The following days the same situation was presented but despite resistance to air for 2 minutes or more on several trials no attacks occurred. Instead the rat showed a new abortive jump which consisted of a little short jump which was just sufficient to remove it from the stand so that it fell directly to the net.

During the last 18 days of testing the screen above the cards was always present and much of the time screens were placed on the sides of the cards to prevent abortive jumping to the right. The one-window situation was presented on 11 test days, the discrimination situation on 3, and the two-negative-card test on 4 days. The rat freely resorted to various types of abortive jumps and when screens were used to block the abortive jumps to the right, it jumped lightly to the ledge and hung there.

On two occasions attacks occurred which were of such a mild nature as to be highly questionable. One of these occurred on the day following a series of tests on the old apparatus in which attempts were made to block abortive jumping by moving the jumping box to the side and so changing the relationship of the jumping box to the cards. The rat adjusted to these, but the following day on the new apparatus it had a slight attack on the first trial. The other mild attack occurred in the one-window situation on the first trial following 3 days of testing on the two-negative-card situation.

During the first 8 of these last 18 test days the routine of abortive behavior was not disturbed by the experimenter and during this period no attacks occurred. This long stretch of testing without the occurrence of an attack rules out the possibility that attacks occur at certain intervals and that the experimenter then rationalizes their cause. It was only during the last 10 days when the situation was altered that behavior approaching the neurotic pattern became evident.

On 5 different occasions during the 58 day period of testing, the front side of the jumping box was closed with a board and Rat 32 was confined in it and subjected to air for a 2 minute period. Although it crouched and clawed at the floor and showed some nervousness nothing approaching an attack oc-

curred on any of these occasions. These tests further substantiate the theory that the attacks in this rat are not direct responses to the air stimulation.

*Summary*

During 58 days of testing with Rat 32 it exhibited what might be called attacks on 14 occasions. The appearance of these attacks becomes clear only when we regard them as the product of the necessity of responding in a situation in which a mode of responding is unavailable. The record of this rat shows that with the appearance of abortive behavior in response to the one-window situation attacks no longer occur. Before cure can be claimed, however, it must be shown that the disappearance of the abnormal pattern is not due to some other cause, such as a change in the physiological make-up of the animal. To rule out this possibility it was shown that (1) the rat could be caused to have attacks in another situation (*e.g.*, the two-negative-card test in the old apparatus) in which abortive behavior was not shown and (2) attacks in the situation in which abortive behavior appeared could be reinstated when the abortive behavior was eliminated.

After the rat developed a cure for the one-window situation a similar cure for the two-negative-card situation was also developed. Finally, it was practically impossible to elicit an attack apparently because a variety of forms of abortive behavior became available to the rat.

To rule out the possibility that attacks occurred only at certain intervals it was shown that the rat could be tested over an 8 day period of time without having an attack. This period is much longer than any test period given this rat without its showing an attack.

In the study of this rat the possibility that exposure to air might be wholly responsible for the attacks was also tested. It was thought, for example, that conflict situations might increase resistance to jumping and hence increase the exposure to air. If this were true the association between conflict and attacks might be purely an artifact. By keeping the intensity of air constant and measuring time it was possible to test this possibility. The records show that when attacks occur

the exposure to air is actually less. In other words, it takes more air to cause the rat to jump than to have an attack. The negative results obtained when the rat was confined in the box and exposed to air but not confronted with stimulus cards demonstrate the function of the stimulus cards in producing attacks.

### *Rat 34M*

*Previous History*

On the old apparatus 42 percent of 12 test days yielded attacks. The most effective situation for this rat seemed to be the presentation of two identical gray cards. After more than a month of no testing this rat was transferred to the new apparatus.

*Critical Experiments and Cure*

During the first 8 days attacks were obtained on the first and third test days. The first occurred on the first presentation of the two identical gray cards although 5 trials on the discrimination problem and 5 trials on the one-window situation failed to produce attacks. On these trials the animal leaped directly to the net or showed a right abortive jump in which it went completely beyond the test cards. The second attack occurred on the third day when the one-window situation was presented. No attacks occurred on the next 5 days but 'escape' reactions (climbing to the top of the jumping box) became more and more regular modes of behaving. It appeared that the rat no longer reacted to the cards but instead the air seemed to produce a fixed escape pattern.

The windows were then left open (no cards in place) and the rat was induced to jump through them. Cards were then placed in the windows but they were tilted backward at about a 30° angle. The windows were thus partly open and the rat could jump onto a card. The cards were then raised more and more from trial to trial and soon the rat jumped against the cards again. This procedure was followed for two days.

The next day (11th) the two identical gray cards were presented at the outset and an attack occurred on the first trial. The following two test days revealed the reappearance

of the escape behavior. Failing to obtain an attack on the second of these days the open-window procedure was again used. A week's vacation was then given to the rat so that the 'escape' behavior might become less available.

An attack occurred the first day after the vacation (14th test day), but no further attacks occurred on the next four test days despite the fact that barriers were introduced on the last few trials of three of these days which were intended to block escape to the top of the box. Although the barriers were used 'escape' responses were attempted and these were either successful or caused the rat to fall to the net. Thus the barriers failed to eliminate the tendency to show the 'escape' response.

On the last of these four test days and on the following day the open-window procedure was again used. At the end of the second day (19th test day) cards were inserted and the rat jumped through the unlocked cards. The cards were then locked. On the first trial the rat jumped to the right card; on the second, to the left card; and on the third, an attack occurred, although the time (20 seconds) before having the attack was no greater than the time preceding a response on the previous trials.

The use of barriers gradually reduced the rat's attempts to climb to the top of the jumping box but instead it jumped abortively to the right of the cards. When barriers were used to block the right abortive jumps the abortive behavior was modified to meet the situation (*e.g.*, jumping between barrier and screen, below the barrier, climbing on barrier, etc.). For two days barriers were used to interfere with the rat's responses but attacks did not occur. On the first of these days the two gray cards were presented side by side and on the second one gray card was presented in the one-window situation.

Failing to obtain an attack the card was removed from the window and the rat was again induced to jump through the open window (22nd test day). As soon as the rat responded readily the card was replaced but left unlatched. After 55 seconds' hesitation the rat jumped to the card. On the next

trial it jumped at the card after 30 seconds and had an attack after striking the card.

On the following three days the rat refused to jump to the cards, but abortively jumped to the right of the cards or climbed out of the box ('escape' pattern). During the next three days the rat was trained to form a discrimination between two cards with opposite figure-ground relations (the training pair). The training procedure was the same as used for an inexperienced rat. On the last of these days (28th test day) a definite preference was shown for the positive card. The two gray cards were then presented and the rat had an attack after 25 seconds in the situation. Since the discrimination behavior on this and previous days was preceded by an average of 40 seconds of hesitation the length of exposure to the situation cannot be held responsible for the attack.

During the next five days the gray cards were presented in the one- and in the two-window situations. 'Escape' behavior was blocked by a screen and abortive jumping to the right and sometimes above the cards appeared regularly after an average resistance of 59 seconds. One attack occurred during this period and it occurred on the fourth trial of the fourth (32nd test day) of these five days. The resistance preceding the attack was 25 seconds, or less than half the average of the other trials.

Since the same situation may sometimes produce attacks and at other times fail to do so, one may suppose that the attack depends upon the physiological condition of the animal rather than the situation. The occurrence of an attack may thus be due to a recurring physiological condition. Thus far the longest interval between attacks occurred after the first two attacks when the abortive behavior was not disturbed. The interval between the second and third attacks was 7 days and the average interval between attacks was 3.4 days. We therefore decided to test the animal on 10 occasions without interfering with the abortive behavior. During these 10 days no attacks occurred and escape behavior became the dominant response. On each of these days resistance be-

havior occurred which exceeded 30 seconds and resistance periods of 1 minute were not uncommon. Since the longest period preceding an attack was 25 seconds in this rat, one cannot claim that the rat reacted before the critical exposure time was reached.

During the next 6 days the rat was again trained (by means of the open-window procedure) to jump to the cards. This relatively long period was necessitated by the fact that abortive jumping occurred as soon as a card was placed in the window. By having the windows half open and by lowering the card after a period of resistance the rat began to respond to the cards. Exposures to the situation exceeded 1 minute on each of these days.

The next day (50th test day) the card was in the window at the outset. After over 70 seconds' resistance the card was lowered and the animal jumped. On the next trial the window was again closed and an attack occurred after 20 seconds.

Tests were continued for 10 more days, the windows being open and closed on various trials but no further attacks occurred.

In the sense that this rat had no attacks in the situation unless its abortive behavior was affected it may be said that the animal was cured.

*Summary*

During the present period of testing Rat 34 had 9 attacks, seven of which occurred after the animal appeared to have become adjusted to the new situation. Its mode of adjustment seemed to be refusal to react to the cards and instead it climbed out and to the top of the jumping box. Blocking this mode of escape resulted in the appearance of an abortive jump to the right. In order to cause the rat to react to the cards retraining was instituted. Six of the seven attacks followed such periods of retraining and in all cases these attacks occurred immediately after the critical situation was reintroduced. After such attacks the rat returned to abortive forms of behavior.

Since the exposure to air in the situation was less before an attack than on control test days one cannot argue that the

manipulation of the situation increased the delay in response and thus increased the period of confinement in the jumping box.

The possible influence of a recurring physiological condition in the animal on its susceptibility was controlled by effecting a period of 10 test days without an attack.

### *Rat 20M*

*Previous History*

This rat developed its neurotic behavior in the discrimination situation. It had developed a right position habit (from which it did not deviate in 200 trials) and all of its attacks occurred when the negative card was on the side of the position habit. Maier (1) interpreted the conflict as similar to that arising in the one-window situation since this animal's position habit forced it to react to but one of the cards. The frequency of attacks for this animal during the last 10 test periods was 40 percent.

*Tests on the New Apparatus*

When tests on the new apparatus were made on this animal the position habit seemed no longer to dominate its behavior. Only 1 error was made in 5 trials despite the fact that 3 of the trials required a response opposite the position preference. Since a differential reaction to the positive and negative cards was shown this animal was now tested in the one-window situation.

During the first 12 test periods (distributed over a period of 17 days) attacks were obtained on 8 days (66.7 percent of the time). The negative card was presented first on 8 of these test periods and attacks occurred on 7 of these days, all of which occurred on the first trial. The positive card was presented first on the other 4 of the 12 test periods. On 3 of these no attack occurred and on the other occasion an attack occurred on the 6th trial when the negative card was in place and after a rest period had been given following the 5 preceding trials. Thus all attacks occurred when the animal was required to jump to the negative card.

Unlike Rats 32 and 34, this rat's attacks have actually increased rather than decreased in the new test situation and correlated with this fact is the fact that this animal failed to show abortive behavior. On days in which no attacks were obtained the animal leaped directly at the cards and on 4 days in which an attack occurred it followed a normal jump to the negative card and the resultant fall to the net. On such occasions the rat remained quietly in the net for several seconds (the air stimulation having been withdrawn as soon as the jump occurred) before it had its attacks.

The last 3 attacks, however, were not accompanied by a leap to the card. In two of them the attack followed a leap to the net and on the last it followed an abortive jump to the right side of the card. Thus some sign of abortive behavior was becoming apparent.

After an interval of a week a second period of testing was given. On some occasions the discrimination problem was presented and on others the one-window situation was used. During this time abortive behavior definitely became evident. Three kinds of abortive responses frequently appeared in the one-window situation. These were progressively as follows: (1) hanging to the ledge; (2) leaps to the left side of the card, and (3) 'escape' responses.

On the 10 occasions in which the one-window situation was presented, attacks occurred on only 2 occasions and these occurred on the first trial. This marked reduction in attacks is thus associated with the increased tendency to show abortive behavior and occurs despite the fact that the continuity of the 10 tests was twice interrupted by several days of testing on the discrimination problem. (Changing the test procedure tends to interfere with adjustments and so increases the attacks.)

The discrimination problem was presented on 18 occasions, 6 of which were interchanged with the one-window tests and 12 of which followed the completion of the one-window tests.

The discrimination tests immediately revealed a strong left position preference. During the first 4 test periods only one response was made to the right. That the animal actu-

ally differentiated the cards was apparent, however, from the fact that it resisted jumping whenever the position habit forced it to leap to the negative card. Thus a position fixation opposite the one formed on the old apparatus had become apparent. Due to the influence of tests in the one-window situation the position fixation in the discrimination problem became replaced by escape behavior. At first the escape response was shown only when the negative card was on the 'position-habit' side but soon the escape behavior was shown for both cards.

During the 18 periods of testing on the discrimination problem only 4 attacks were obtained and all of them occurred when the negative card was on the side of the position preference. Two of them occurred during the period interrupted by tests on the one-window situation and 2 during the last 12 days in which the procedure was not altered.

Since attacks occasionally occur in this rat a complete cure cannot be claimed.

*Summary*

The record of Rat 20 demonstrates a relationship between abortive behavior and attack frequency. Its record also furnishes evidence showing that the attack is partly a function of the reaction to the cards rather than purely a direct response to the air. This is shown by the fact that all of this rat's attacks in the one-window situation occurred in response to the negative card and all of its attacks in the discrimination situation were in response to the negative card when it conflicted with the position habit.

### *Rats 12M and 25F*

In tests in the one-window situation Rat 12 had attacks on 7 out of the 12 tests on the old apparatus. On the new apparatus, attacks were obtained on the 1st trial of the 1st and 2nd days, but on the following 5 days no attacks occurred and the animal showed left abortive responses. During the next 6 days, tests on the old and new apparatus were alternated. Three tests on the old apparatus yielded 2 attacks and 3 tests on the new apparatus yielded none.

The rat was then subjected to cortical operation so that further data bearing on the present problem were not obtained.

Rat 25 had attacks in 3 out of the 11 last systematic tests on the old apparatus. In 10 tests on the new apparatus it had attacks on the 1st trial of every test. Because the attacks occurred on the 1st trial abortive behavior could not be studied. Since the attacks of Rat 25 increased on the new apparatus it can hardly be clained that the air was less effective in the new apparatus.

### *Rat 75M*

*History*

Rat 75 had previously learned the discrimination problem (black circle on white vs. white circle on black background) and although air was used to hasten its responses no attacks were obtained. To determine its susceptibility to purely auditory stimuli it was then placed in the net of the jumping apparatus and exposed (for 2 minutes) to intense sounds of rushing air. No attacks to air were obtained in 10 such tests. The jingling of keys was then substituted for the sound of air and one attack was obtained in 10 such exposures.

*The Manipulation of Test Conditions in the Old Apparatus*

After an interval of several weeks retests on the discrimination problem were given and perfect retention of the habit was shown.

The one-window situation was then introduced. An attack was obtained on the first trial when the rat was required to jump to the negative card. No further attacks occurred during the next 6 days of testing in this situation. On these trials the rat jumped abortively to the right of the card whenever the negative card was presented.

A screen was then introduced which prevented the abortive jumping to the right. This screen was used during the next 9 days and 4 attacks were obtained, all of which occurred during the first 5 days and when the rat was required to jump

to the negative card. When no attacks occurred the jumps were directed at the card.

Resistance to jumping when the negative card was in place was greater on trials not followed by attacks than on trials when attacks occurred. The average resistance of jumps followed by attacks was 95 seconds and for those followed by a normal jump, 175 seconds.[3]

Following the last 4 days of testing with the screen placed to the right 5 more days of testing without the screen were given. During the first 2 days the rat continued to jump directly at the cards, and thereafter abortive jumps to the right occurred whenever the negative card was in place. No attacks were obtained on these days.

*Summary*

The history of Rat 75 showed it to be unaffected by air stimulation until it was placed in the one-window situation for the first time. It then seemed to develop an adjustment through abortive jumping. This interpretation is supported by the fact that the interference with the abortive jumping (by the addition of a screen) resulted in a reinstatement of the attacks, 4 of which occurred during the first 5 days in which the screen was used. That attacks did not continue suggests that a further adjustment of some kind was developed.

### *Rat 66M*

Three rats (58M, 64M, and 66M) which failed entirely to have an attack through the usual procedures in the laboratory were subjected to a strenuous period of testing including all forms of auditory stimuli, enclosed box with air, one-window situation, and two-negative-card procedure but no attacks were obtained. All showed considerable facility with abortive forms of jumping.

The present experimenter then tested the rats on the new apparatus but with the screen above the cards as on the old apparatus. On the first day (5th trial) Rat 66 had a violent

[3] In the case of this rat the air was applied in steps as follows: 30 sec., no air; 30 sec., mild air; 30 sec., moderate air; thereafter, full air intensity. This rat's resistance should therefore not be compared with that of other rats reported in this study.

attack. Rat 64 refused to jump (but did so the following day) and Rat 58 jumped abortively.

Rat 66 was tested on 13 occasions during the next month and 2 further attacks were obtained but both were definitely less intense than the first and convulsions were lacking.

Of particular interest is the fact that this rat became extremely retiring in its cage behavior after the first attack. This retiring behavior continued not only during the following month when tests were periodically given but lasted until its death several months later at the age of 2 years.

The case of Rat 66 is of interest because its behavior remained stable until it was transferred to a different apparatus. Although a discrimination habit may be successfully transferred when the apparatus is changed, the manner of jumping (in terms of resistance and abortive behavior) is frequently greatly altered. It often occurs that different abortive responses are shown on two apparatuses when these are alternated. All three of the rats mentioned above transferred their discrimination habit to the new apparatus, but the abortive behavior was modified in each case. It seems reasonable to conclude that Rat 66's attacks occurred because the change in apparatuses interfered with its abortive responses.

### *Rats 61M, 62M, and 63M*

Rats 61, 62, and 63 had attacks when subjected to the hiss of air and the jingling of keys as well as when placed in the one-window situation. Ten exposures to each of the above tests showed the following results. Rat 61: 2 attacks to air, 4 to keys, and 9 to the one-window situation. Rat 62: 2 attacks to air, 2 to keys, and 7 to the one-window situation. Rat 63: 3 attacks to air, 9 to keys, and 9 to the one-window situation.

On the new apparatus Rats 61 and 62 were caused to jump over the stimulus cards whereas Rat 63 jumped abortively to the left. The procedure was to apply air for short intervals at a time. With this manner of jumping the rats were caused to jump on ten successive days (5 to 10 trials per day) with-

out having attacks. Previously the animals could not be caused to leave the starting box without having an attack.

Attempts were then made to apply air persistently and intensively. Under these conditions attacks again occurred and in no case can anything like a cure be claimed. The best one can say is that the abortive behavior made it possible for the animals to escape air stimulation by jumping more readily. When, however, escape was delayed, either by the experimenter or because of the rat's hesitancy, attacks frequently occurred.

Whether or not more patient handling would have produced positive results cannot be said, but certainly animals which have attacks to air stimulation alone present a picture quite different from those discussed above.

## Summary and Conclusions

The purpose of this experiment was to test the hypothesis that abortive behavior may be an alternative to the neurotic pattern. Attempts were therefore made to manipulate the animal's response tendencies by modifying the jumping apparatus so that the same problem could be given under conditions which differed in the degree to which abortive behavior was possible. It was found that there was a close relationship between the appearance of abortive behavior and the disappearance of neurotic pattern in seven rats. In two of the cases the experiment was continued to the point where the neurotic pattern was eliminated and reinstated several times by controlling the abortive response. Where such control was demonstrated it was believed the animal could be said to be cured by the experimenter. In all seven of these rats the reaction to the test cards was shown to play a fundamental part in the production of the neurotic pattern.

In three other cases in which air stimulation alone was adequate to produce attacks (although the jumping situation increased their frequency) attempts to produce a cure failed. However, it can be said that the development of abortive responses decreased the resistance to jumping and in this manner the animals were able to avoid attacks.

Because abortive behavior may be said to resolve the conflict created by the excitation produced by air and the inhibition produced by the negative member of the discrimination cards, the presence of conflict seems to be a basic factor in the appearance of attacks. If conflict were not important it is difficult to see how abortive behavior could be an alternative to the neurotic pattern. To show that some rats show attacks to air stimulation alone does not rule against this interpretation. Rather it may be said that in such animals conflict is either not essential or that it is present in a less obvious form. Maier and Glaser (**2**) have presented evidence to support the latter alternative.

(Manuscript received April 27, 1940)

## References

1. Maier, N. R. F., *Studies of Abnormal Behavior in the Rat.* I. The neurotic pattern and an analysis of the situation which produces it. Harper and Bros., New York, 1939, pp. 81.
2. Maier, N. R. F., and Glaser, N. M., Studies of abnormal behavior in the rat. II. A comparison of some convulsion-producing situations, *Comp. Psychol. Mono.*, 1940, 16, pp. 30.
3. Maier, N. R. F., Glaser, N. M., and Klee, J. B., Studies of abnormal behavior in the rat. III. The development of behavior fixations through frustration, *J. exp. Psychol.*, 1940, 26, 521–546.
4. Maier, N. R. F., Abortive Behavior as an Alternative for the Neurotic Attack in the Rat. Film. Psychological Cinema Register. Editor, A. Ford, Bethlehem, 1940, 600 ft. 16 mm.

# AN EXPERIMENTAL INVESTIGATION OF TWO FACTORS WHICH PRODUCE STEREOTYPED BEHAVIOR IN PROBLEM SITUATIONS

BY ELSA E. ROBINSON

*New York University*

An observation frequently made by experimenters in the field of problem solving concerns the high degree of persistence often shown by subjects in pursuing an unsuccessful mode of activity. Such a problem-solving 'blind spot' entails repetitions of the same inadequate behavior as determined by the same erroneous set or hypothesis. One might think that the encountering of an obstacle to progress should result in the abandonment of the obstructed mode of activity and the initiation of another. Nevertheless, although blocked, the individual often goes over the same ground time after time in the same way, to be brought up short repeatedly at the same point by the same impasse. In fact, as Ruger, Bulbrook, Maier, and others have reported, subjects whose attempts to solve a given problem endured for a considerable period of time made all their responses within the frame imposed by a single original misconception of the task.[1]

Moreover, persistence in a single fruitless plan of attack has been considered a primary obstacle to the solution of problems. Thus Maier has devoted an experiment to the purpose of discovering whether the erroneous set is a factor which actually inhibits the emergence of other sets, or whether it is merely "a function which is present during the absence of more fruitful processes." He concludes that the former is the case.[2] Several experimenters have noted that stimuli which might have provided the cue for the correct mode of

[1] H. A. Ruger, The psychology of efficiency, *Arch. Psychol.*, 1910, 2, No. 15, 1–88; M. E. Bulbrook, An experimental inquiry into the existence and nature of 'Insight,' *Amer. J. Psychol.*, 1932, 44, 409–453; N. R. F. Maier, An aspect of human reasoning, *Brit. J. Psychol.*, 1933, 24, 144–155.

[2] *Op. cit.*

attack were either not responded to at all, or responded to in terms of the prevailing incorrect set. Thus Ruger writes:[3]

> In some cases the appropriate variation would take place many times but would pass unnoticed. The assumptions often had an inertia, an apparently volitional persistency about them. . . . The essential parts of the apparatus might be selected but with the wrong assumption. This assumption, then, either limited the motor variations or made the subject insensitive to variations not in line with the assumption in question.

Bulbrook, discussing the behavior of subjects engaged in solving a mechanical puzzle, writes:[4]

> . . . the first self-instruction of manual dexterity was never changed—a self-instruction inhibiting the apprehension of relevant conditions which would necessitate a reformulation of the plan of procedure.

And Maier:[5]

> We often miss obvious solutions because we try to solve problems in an habitual way (which is often the wrong 'direction') and are thus blind to the obvious solution. We misinterpret the ideas of others to fit our own 'direction' in the same way that suggestions in this study were made to fit the 'direction' that the subject already had.

And Maier again:[6]

> . . . 'directions' were so decided and dominating that they caused helpful suggestions to be so applied as to fit in with the organizations of experience which they were determining. Consequently many suggestions were useless and even misleading.

It is obvious that the longer the incorrect set persists the longer is the solution delayed. One of the most important tasks in the investigation of problem-solving behavior, therefore, is the discovery of the factors which determine the duration of behavior which fails to lead to a desired goal. Such behavior must be conditioned upon experimentally ascertainable factors. Two of these factors are studied in the experiments reported below. In the first experiment, the variable investigated is the *degree of the subject's confidence* that the activity in which he is engaged will result in the attainment of the end which he seeks. In the second experiment, the variable is the *number of alternative activities of equal*

[3] *Op. cit.*, p. 17.

[4] *Op. cit.*, p. 439.

[5] Reasoning in humans. I. On direction, *J. comp. Psychol.*, 1930, **10**, 140–141.

[6] Reasoning and learning, *Psychol. Rev.*, 1931, **38**, 338.

*promise* which the subject is prepared to engage in should his first mode of behavior fail to achieve his goal.

## Experiment I

### *Purpose*

The aim of this study was to discover whether any relation obtains between the individual's perseverance in a given activity and the degree to which he feels assured that further pursuit of that activity will result in the attainment of his end.

### *Subjects, Materials, Procedure*

*Subjects.*—Fifty-two college students participated in the experiment.

*Materials.*—The problems presented were anagrams. For the specific factors to be controlled in this experiment, it was believed that verbal material of the anagram variety possessed advantages over other kinds of test situation. In order to test the effects of the experimental variable, *i.e.*, the degree of the subject's confidence in his ability to solve concrete problems, it was necessary to present a number of problems of the same 'genre.' Furthermore, these problems had to be as nearly as possible equal for the subject both in interest and in apparent difficulty. Finally, in order to measure the factor of perseverance, it was of course necessary that the problems be insoluble. Obviously, too, it was necessary that the subject have no suspicion that the problems were 'unfair' in this sense. Most insoluble problems of a mathematical or parlor-game variety would sooner or later be detected as insoluble, if not by the majority of subjects, at any rate by the brightest, or by those most familiar with the given concrete form of problem. The possible combinations of anagram lay-outs of the kind used (10 letters) can be completely tested only after very protracted labor, no matter how intelligent the subject or how great his aptitude for anagrams.

Eight anagram lay-outs were constructed. Each was ten letters in length, with no letter repeated. Each was composed of eight consonants and two vowels. Every lay-out was printed in identical form on ten cards. Eighty cards thus constituted the complete set. A five-letter or longer word was the criterion of a solution. Four of the lay-outs were easily solvable. They were used for purposes of general motivation and to confirm the subject's trust in the veracity of the instructions. These anagrams were:

(1) U M S N D B H O R T
(2) S P I L N M G Y T A
(3) L E P C T I M D R V
(4) A T L R F V S N K E

The following lay-outs constituted the test problems. None of the four can be solved for a 5-letter word.

1. E X R V L Z W U P F
2. J F M A Z T K B R U
3. C F E H M L G Z X O
4. N T Z V O D W A J C

*Procedure.*—The following general instructions on the method of playing the game of anagrams were read:

> This is an experment aimed at discovering how people solve anagrams. You know what an anagram is, don't you? You are given certain letters in a jumbled

order, and your task is to put them into a form which gives you a correctly spelled word. Thus, if you were given these letters [E presented a card with the printed inscription T F E A N R K I G O], you could form a good many words. Try to find some 5-letter or longer words in the jumble. [If S was not facile at this, E pointed out several possibilities.]

Certain rules are to be observed in playing the game:

1. Any English word is acceptable. Foreign words are not.
2. Proper nouns, such as the names of people and places, are not acceptable.
3. You may not use a letter in a word twice unless it occurs twice in the letter jumble.
4. Any word is precisely as good as any other word. If you discover a common word announce it immediately without stopping to discover a more unusual word. What I want is your first response. I am not at all interested in the rareness or the 'literary value' of the word.
5. Any inflected form of a word is all right. 'Frozen' is as acceptable as 'freeze'; 'women' is just as good as 'woman.'
6. You will not be marked for the length of time you take to arrive at a solution. It follows that you are not to work with the idea that the faster you solve an anagram, the higher your record. When you are given a letter jumble, you are immediately to set about trying to form a word. Make every effort to work conscientiously and without interruption. But I have no interest in the speed of your performance.

These preliminary directions were followed by the instructions for the experiment proper.

I am going to hold 10 cards in my hand, face down. On each card there is a set of letters, and each card contains a different set of letters. You are to select one card, place it before you, and try to get a word which is at least 5 letters long. Longer words are all right, but they are no better than 5-letter words. You will get one and only one card of the 10 I hold out to you. Sometimes all 10 cards can be solved, that is to say, they will each contain at least one 5-letter word. Then you will know that whatever card you choose, a solution is possible, although you may not be able to find it. Sometimes, only 9 cards will contain letters which can be put together to form a word. In this case, the tenth card will consist of letters which cannot be solved at all, no matter how hard or how long you try. Since you will choose a card blindly, you will know that with that card, you have 9 chances out of 10 that you can get a word, even if you are a wizard at anagrams. Sometimes, of the 10 cards, only 5 will contain possible solutions; then you will have only 5 chances out of 10 of having a solvable problem before you. Sometimes, the cards will be heavily stacked against your luck; only 1 of the 10 cards from which you choose will contain an anagram which can be solved. And there will be other arrangements of chances.

Each time I hold out a handful of 10 cards, I will tell you how many of them contain anagrams which can be solved. The remainder cannot be solved by anyone; there isn't any solution at all. So that you can figure out your chances of picking a possible winner.

Remember, you are to select a card from the handful I hold out to you, and try from the letters on it to form an English word of 5 letters or more. As soon as you have found one, announce it. Do not pay attention to or announce words

of 4 letters or less. I am not interested in the length of time you take studying each card. When and if you are reasonably sure that you cannot solve the card before you, turn it to the blank side.

Four degrees of subjective assurance of the ability to solve were tested. Obviously these 'degrees of self-confidence' have no absolute meaning for a single subject, much less a uniform meaning for different subjects. It was thought, however, that different feelings of solution-possibility should be induced when the experimenter declared that the chances on a given problem were perfect, just short of perfect (9 chances out of 10), equalled the chances that the problem was insoluble (5 out of 10), or that the chances were greatly against the subject (1 out of 10). While it was surely not assumed that such instructions elicited certain fixed 'amounts' of a confidence attitude from the subject, it did not seem hazardous to predict that the subject's relative degree of confidence in his ability to solve a problem would vary roughly with his 'knowledge' of its solution-possibility.

In presenting a given problem, the experimenter always held the ten cards composing that set spread out fan-wise toward the subject. The subject was told how many of the ten cards were solvable and how many were not. Actually, since all ten cards composing each set contained the same letter lay-out, and since for the four test sets no solution was possible, the subject was being deliberately misinformed. Then the subject was permitted to draw one card. For the four test anagrams the time was recorded from the moment the subject started to study the card till he reversed it. The time was kept in minutes and seconds by means of an electric clock which was invisible to the subject. No time record was kept for the solvable problems, all four of which were solved by all subjects usually after a few seconds of scrutiny. The experimenter sat opposite the subject and pretended to be engaged in the reading of a novel.

The order of presentation of the anagrams was always the same: (1), 1, (2), 2, 3, (3), 4, (4). Or in terms of the possibility of solution: S (solvable), I (insoluble), S, I, I, S, I, S.

A group control was devised in order (1) to provide a control of the fatigue and practise factors, and (2) to constitute a check against the possibility of any general differences in the 'apparent difficulty' of the four test problems. The subjects were divided into four groups, each consisting of 13. For each of these groups the presentation order of the 8 problems was uniform (the order indicated above). The chances assigned to the four *solvable* problems were constant for the 4 groups: 7 out of 10 chances for Anagram (1), 10 out of 10 for (2), 6 for (3), and 9 for (4). But the order of designated chances for the four *test* problems differed for the various groups, as follows:

Number of Chances out of a Total of Ten

| Anagram | Group I | Group II | Group III | Group IV |
|---|---|---|---|---|
| 1 | 10 | 5 | 1 | 9 |
| 2 | 9 | 10 | 5 | 1 |
| 3 | 5 | 1 | 9 | 10 |
| 4 | 1 | 9 | 10 | 5 |

Thus, for example, the 10 cards of Anagram 1 were presented to Group I as all solvable, to Group II as 5 solvable and 5 not, to Group III as all insoluble except for one card, and to Group IV as all solvable except for one.

## *Results*

Table I presents the complete records for each of the 52 subjects.

TABLE I

INDIVIDUAL RECORDS OF 52 SUBJECTS:
TIME SPENT BEFORE ABANDONING EACH OF FOUR PROBLEMS WHICH WERE ALLEGED TO DIFFER IN THEIR CHANCES OF SOLUTION

Group I

| Subject | Number of Chances in 10 | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 10 | | 9 | | 5 | | 1 | |
| 1 | 17′ 20″ | 6.01 | 6′ 52″ | 2.38 | 4′ 54″ | 1.58 | 2′ 53″ | 1.00 |
| 2 | 13′ 43″ | 4.52 | 5′ 58″ | 1.97 | 4′ 10″ | 1.37 | 3′ 2″ | 1.00 |
| 3 | 8′ 14″ | 2.41 | 8′ 40″ | 2.54 | 5′ 48″ | 1.70 | 3′ 25″ | 1.00 |
| 4 | 2′ 50″ | 4.47 | 1′ 30″ | 2.27 | 1′ 20″ | 2.11 | 38″ | 1.00 |
| 5 | 21′ 7″ | 5.44 | 9′ 22″ | 2.41 | 5′ 48″ | 1.49 | 3′ 53″ | 1.00 |
| 6 | 7′ 55″ | 6.51 | 3′ 34″ | 2.92 | 2′ 25″ | 1.99 | 1′ 13″ | 1.00 |
| 7 | 2′ 50″ | 2.62 | 2′ 15″ | 2.08 | 1′ | .92 | 1′ 5″ | 1.00 |
| 8 | 3′ 2″ | 3.57 | 1′ 33″ | 1.82 | 1′ 39″ | 1.94 | 51″ | 1.00 |
| 9 | 16′ | 3.10 | 10′ 20″ | 2.00 | 6′ 22″ | 1.23 | 5′ 10″ | 1.00 |
| 10 | 19′ 40″ | 5.54 | 15′ | 4.23 | 8′ 56″ | 2.52 | 3′ 33″ | 1.00 |
| 11 | 6′ 22″ | 14.15 | 2′ 5″ | 4.63 | 1′ | 2.22 | 27″ | 1.00 |
| 12 | 18′ 22″ | 4.32 | 11′ 48″ | 2.78 | 3′ | .71 | 4′ 15″ | 1.00 |
| 13 | 6′ 10″ | 3.08 | 3′ 5″ | 1.54 | 1′ 44″ | .87 | 2′ | 1.00 |

Group II

| Subject | Number of Chances in 10 | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 5 | | 10 | | 1 | | 9 | |
| 14 | 2′ 19″ | 4.63 | 6′ 2″ | 12.07 | 30″ | 1.00 | 4′ 15″ | 8.50 |
| 15 | 6′ 5″ | 1.24 | 13′ 33″ | 2.76 | 4′ 55″ | 1.00 | 11′ 16″ | 2.29 |
| 16 | 2′ 17″ | .58 | 20′ 20″ | 5.19 | 3′ 55″ | 1.00 | 10′ 46″ | 2.75 |
| 17 | 3′ 15″ | 2.67 | 17′ 50″ | 14.66 | 1′ 13″ | 1.00 | 7′ 12″ | 5.92 |
| 18 | 1′ 48″ | 1.44 | 12′ 14″ | 9.79 | 1′ 15″ | 1.00 | 7′ 8″ | 5.71 |
| 19 | 2′ 6″ | 1.97 | 8′ 4″ | 7.56 | 1′ 4″ | 1.00 | 5′ 43″ | 5.36 |
| 20 | 2′ 58″ | 1.71 | 6′ 17″ | 3.62 | 1′ 44″ | 1.00 | 5′ 25″ | 3.13 |
| 21 | 3′ | 1.64 | 11′ 2″ | 6.02 | 1′ 50″ | 1.00 | 9′ 53″ | 5.39 |
| 22 | 4′ 15″ | 1.88 | 9′ 31″ | 4.20 | 2′ 16″ | 1.00 | 6′ 52″ | 3.03 |
| 23 | 7′ 12″ | 1.20 | 8′ 13″ | 1.37 | 6′ | 1.00 | 8′ 13″ | 1.37 |
| 24 | 3′ 35″ | 1.82 | 5′ 22″ | 2.73 | 1′ 58″ | 1.00 | 4′ 4″ | 2.07 |
| 25 | 1′ 25″ | .89 | 25′ 28″ | 16.08 | 1′ 35″ | 1.00 | 5′ 45″ | 3.63 |
| 26 | 10′ 27″ | 1.56 | 19′ 4″ | 2.85 | 6′ 42″ | 1.00 | 16′ 15″ | 2.43 |

TABLE I—(*Continued*)

Group III

| Subject | Number of Chances in 10 | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 1 | | 5 | | 9 | | 10 | |
| 27 | 1′ 54″ | 1.00 | 3′ 2″ | 1.60 | 5′ 1″ | 2.64 | 13′ 12″ | 6.95 |
| 28 | 2′ 57″ | 1.00 | 6′ 57″ | 2.36 | 10′ 53″ | 3.69 | 11′ 25″ | 3.87 |
| 29 | 1′ 10″ | 1.00 | 1′ 34″ | 1.34 | 5′ 4″ | 4.34 | 8′ 21″ | 7.16 |
| 30 | 35″ | 1.00 | 53″ | 1.51 | 1′ 45″ | 3.00 | 4′ 48″ | 8.23 |
| 31 | 2′ 8″ | 1.00 | 2′ 20″ | 1.09 | 4′ 28″ | 2.10 | 13′ 31″ | 6.34 |
| 32 | 3′ 57″ | 1.00 | 4′ 30″ | 1.14 | 7′ 52″ | 1.99 | 14′ 17″ | 3.61 |
| 33 | 1′ 20″ | 1.00 | 2′ 37″ | 1.96 | 3′ 14″ | 2.43 | 7′ 28″ | 5.60 |
| 34 | 55″ | 1.00 | 1′ 58″ | 2.15 | 6′ 22″ | 6.95 | 12′ 52″ | 14.04 |
| 35 | 2′ | 1.00 | 3′ 10″ | 1.58 | 5′ 14″ | 2.62 | 8′ 17″ | 4.14 |
| 36 | 5′ 8″ | 1.00 | 7′ 52″ | 1.53 | 14′ 19″ | 2.79 | 23′ | 4.48 |
| 37 | 1′ 22″ | 1.00 | 2′ 24″ | 1.76 | 7′ 46″ | 5.68 | 18′ 45″ | 13.72 |
| 38 | 15″ | 1.00 | 25″ | 1.67 | 36″ | 2.40 | 1′ 20″ | 5.33 |
| 39 | 2′ 39″ | 1.00 | 3′ 15″ | 1.23 | 5′ 31″ | 2.08 | 12′ 32″ | 4.73 |

Group IV

| Subject | Number of Chances in 10 | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | 9 | | 1 | | 10 | | 5 | |
| 40 | 5′ 52″ | 1.03 | 5′ 40″ | 1.00 | 14′ 23″ | 2.54 | 3′ 50″ | .68 |
| 41 | 7′ 17″ | 1.77 | 4′ 7″ | 1.00 | 13′ 20″ | 3.24 | 5′ 38″ | 1.37 |
| 42 | 2′ 25″ | 1.53 | 1′ 35″ | 1.00 | 5′ 49″ | 3.67 | 2′ 27″ | 1.55 |
| 43 | 9′ 52″ | 4.03 | 2′ 27″ | 1.00 | 20′ 30″ | 8.36 | 4′ 15″ | 1.73 |
| 44 | 3′ 1″ | 1.77 | 1′ 42″ | 1.00 | 6′ 51″ | 4.03 | 2′ 47″ | 1.64 |
| 45 | 9′ 50″ | 3.51 | 2′ 48″ | 1.00 | 11′ 57″ | 4.27 | 7′ 43″ | 2.76 |
| 46 | 13′ 27″ | 5.64 | 2′ 23″ | 1.00 | 22′ 4″ | 9.26 | 6′ 39″ | 2.79 |
| 47 | 5′ 43″ | 1.06 | 5′ 23″ | 1.00 | 12′ 12″ | 2.27 | 6′ | 1.11 |
| 48 | 5′ 35″ | 1.46 | 3′ 50″ | 1.00 | 11′ 17″ | 2.94 | 4′ 44″ | 1.23 |
| 49 | 6′ 53″ | 1.78 | 3′ 52″ | 1.00 | 9′ | 2.33 | 4′ 20″ | 1.12 |
| 50 | 7′ 42″ | 2.73 | 2′ 49″ | 1.00 | 14′ 10″ | 5.03 | 4′ 53″ | 1.73 |
| 51 | 4′ 19″ | 3.41 | 1′ 16″ | 1.00 | 8′ 53″ | 7.01 | 1′ 12″ | .95 |
| 52 | 7′ 18″ | 5.28 | 1′ 23″ | 1.00 | 13′ 37″ | 9.84 | 2′ 40″ | 1.93 |

The values to the left of each column represent time, in minutes and seconds.

The values to the right of each column represent time-ratios, relative to condition '1 chance in 10' taken as unity.

Two deductions are immediately apparent from a study of the time data.

1. There is very little consistency from subject to subject as regards the absolute time spent on a problem having a designated number of chances of solution. The time ranges for the four conditions are:

10 chances in 10: 1′ 20″ — 25′ 28″
9 chances in 10: 36″ — 16′ 15″
5 chances in 10: 25″ — 10′ 27″
1 chance in 10: 15″ — 6′ 42″

2. There is a definite and remarkably consistent tendency for the time spent before resigning a problem to vary directly with the designated chances that the problem is a solvable one.

To facilitate such comparisons, all the time values were transmuted into terms of ratio. This was done in the case of each subject by making his time value for the '1 chance in 10' anagram equivalent to 1.00, and by expressing the time values for the three other conditions of solution-likelihood (tested by the three other anagrams) as ratios of the time for the '1 chance in 10' problem. These ratio values appear on the right side of each column in Table I.

The time-ratio ranges for the three conditions in terms of the '1 chance in 10' condition are as follows:

| | |
|---|---|
| 10 chances in 10: | 1.37 – 16.08 |
| 9 chances in 10: | 1.03 – 8.50 |
| 5 chances in 10: | .58 – 4.63 |

In the attempt to analyze the differences statistically, it was found that the distributions of the 52 *raw time* values for each degree of confidence deviated considerably from the normal. Similarly, the distributions of the *ratio values* for each degree of confidence showed marked skewing. A logarithmic treatment of the ratio data was found to be the soundest statistical procedure.[7]

[7] I am indebted to Professor Thomas N. Jenkins of New York University for his aid in the statistical treatment of these data and also of the data obtained in the following experiment. The validity of the logarithmic analysis of the present data is indicated by the close approximation of the limits of each distribution, as defined by $\pm 3\sigma$ ratio, to the actual range of ratio values for that distribution. The discrepancy between the upper empirical limits of the two ratio distributions ('9 chances out of 10' and '10 chances out of 10') and the upper statistical limits defined by $\pm\ 3\sigma$ ratio is accounted for by the fact that the curves for these two distributions of ratio-values are definitely flat. They approach the abscissae very gradually. The accumulation of more data would undoubtedly have furnished values above 9.00 in the case of the '9 chances' distribution, and above 16.50 in the case of the '10 chances' distribution. A description of the method used in the present analysis is to be found in an article by Professor Jenkins, A short method and tables for the calculation of the average and standard deviation of logarithmic distributions, *Annals of Mathematical Statistics*, 1932, 3, 45–55.

A treatment based on the assumption of a normal distribution of the raw time-values or of the ratio-values yields critical ratios of even greater magnitude than are obtained by the logarithmic analysis.

As calculated by the Jenkins method, the geometric means ($g$) of the time ratios for the three experimental conditions, with the condition '1 chance in 10' treated as 100, are:

5 chances in 10: 153.46
9 chances in 10: 281.52
10 chances in 10: 507.76.

The reliability of the differences between the means was tested by the formula:

$$\sigma_{\text{diff.}} = \sqrt{\frac{\sigma g_1^2}{N} + \frac{\sigma g_2^2}{N}}.$$

As Table II shows, all differences between the means for the four experimental conditions are completely reliable.

TABLE II

DIFFERENCES BETWEEN THE GEOMETRIC MEANS FOR THE FOUR EXPERIMENTAL CONDITIONS, WITH STANDARD DEVIATIONS OF THE DIFFERENCES, AND CRITICAL RATIOS

| Conditions Compared | Difference Between Logs of Geometric Means | S.D.$_{\text{diff.}}$ | Critical Ratio |
|---|---|---|---|
| 1 and 5 chances in 10 | .18600 | .023 | 8.09 |
| 5 and 9 chances in 10 | .26351 | .037 | 7.12 |
| 9 and 10 chances in 10 | .25615 | .044 | 5.82 |

These findings are discussed in connection with the results of Experiment II.

## EXPERIMENT II

### *Purpose*

The aim of this study was to discover whether any relation obtains between the individual's perseverance in a given activity and the number of other activities of equal promise available to him which may achieve the same goal.

### *Subjects, Materials, Procedure*

*Subjects.*—The subjects were 42 college under-graduates.

*Materials.*—Anagrams were again used in this experiment. Forty different anagram cards were prepared, each presenting 6 consonants and 2 vowels. Since it

was considered desirable that the 'apparent' difficulty of the various lay-outs be as nearly equal as possible, all the vowel combinations used contained 'e' and the pair 'u e' was omitted.[8] The forty cards consisted of ten each with the vowel combinations 'a e,' 'e e,' 'i e,' and 'o e.' All forty cards differed in the consonant components.

In addition to the anagram lay-outs, 12 small cards were prepared, each containing a single consonant.

*Procedure.*—After preliminary directions on how to play the game of anagrams and the rules to be observed (as in the preceding experiment), the following instructions were read to the subject:

> You will first be given a small card with a consonant printed on it. The object of the game is to make a word at least 6 letters long, beginning with the consonant. In addition to this consonant, you will have 8 other letters. So that you will have 9 letters in all. Now, I am not going to *give* you the 8 additional letters. You will *select* them. It will be done in this way. I will spread out on the table a number of large cards face-down. On the reverse of each card, there will be 8 letters, and each card will contain 6 consonants and 2 vowels. None of the cards on the table will contain the same set of 8 letters. All the cards are different. You will be selecting your set of letters blindly, of course, since the cards will lie with the letter-side down. Sometimes you will be allowed to pick one card, sometimes two, and sometimes as many as seven cards (that is, 7 sets of 8 letters). When you have made your selection, spread your cards out on the table, face-down, and then turn one of them letter-side up.
>
> Now, since you have picked 8 of your letters by chance, it may happen that you will be able from them to form a 6-letter word beginning with the prescribed initial letter easily, or perhaps only with difficulty, or perhaps not at all, that is, perhaps no 6-letter word can be made from the letters you have selected. If you cannot get a word with the first card, you are to keep the initial letter and turn the card you have been working on face-down. Then turn up another card, in the hope of having better luck to form a word starting with the same letter. And so on.
>
> Remember, sometimes you will be allowed 7 chances, sometimes 2 chances, and sometimes, only one. Your only job is to get a 6-letter or longer word, starting with the consonant I give you. If you are allowed 7 chances, it doesn't make any difference for your score in the game, if it takes you all 7 trials, or 5, or 4, or if you are successful on the very first trial. This is naturally so, because it is mostly a matter of luck, whether you draw letters which will go together to form a word easily, or only with much thought, or maybe not at all, no matter how long or how hard you try. Your performance is not being timed. You will probably be able to do several games before the end of the hour. Here is a score sheet. The numbers on the left designate the games. Each game is represented by a new initial letter, no matter how many chances you are allowed. The blanks next to the numbers are where you are to write in the word you get, if you get one. As soon as you get one, write it in the correct space and return all the cards of that game to me. This means you have succeeded on that game,

[8] It was remarked in preliminary observations that the *apparent* solubility of 8-letter anagrams was greater if one of the two vowels was an 'e' than if neither was. It was further observed that when one vowel was an 'e' and the second a 'u,' the anagram seemed more difficult than if the second vowel were any other one.

and therefore that game is over. Then we will start on a new game, that is, on a new initial consonant. Don't forget that you are simply to try to get one word for each starting letter, and it doesn't make any difference how many trials you use of the total number you are allowed. If you fail to get a word for a given initial-letter, leave that space on the score card blank. We will then start on a new initial letter. The object is to achieve a high proportion of correct responses, rather than a large number of them. Thus, if you get 2 right out of 4 initial letters you have tried, you have done better than if you get 3 right out of 7 games. We will now see if you understand how the game is played.

The experimenter gave the subject the initial-letter card, 'S.' The subject was asked to select 7 from a large number of trial cards. These cards were singly studied until a word which satisfied the requirements was found. Any points of the technique which remained obscure to the subject were cleared up in this preliminary illustration. Then the salient points in the instructions were again emphasized in the following terms:

You are to remember 3 things: (1) That you may take as long as you like studying each card. (2) That you are at complete liberty to use all the trials you are allowed. If you are allowed 2 trials, you get no more credit if you succeed on the first than if it takes you both trials. And, of course, since this is mostly a matter of luck in drawing letters, you may fail on the second trial also. In which case, you will not be allowed further trials with that initial letter. (3) Always remember that the object of the game is to try, if possible, to get a word of 6 letters or more, beginning with the prescribed consonant. When I give you an initial letter for the first time, we are starting a game. You lose the game, if you use all your trials for that letter and cannot find a word. You win, if you find a word in any one of the trials, regardless of which one. That ends that game, and then we start a new one using a new and different initial consonant.

In the test series, the order of presentation of initial consonants in successive games was always C, P, T, B, M, D, R, F, H, G, and L.

The experimenter recorded the length of time which the subject spent on the first card of each game before that card was resigned. Or if the subject was allowed only one card, the experimenter noted the time he spent on that one before giving up.

The 42 subjects were each tested in 6 games, 2 each for the conditions 'no alternative,' '1 alternative,' and '6 alternatives.'[9] For each of the following rotational orders, 7 subjects were tested (the numbers 0, 1, and 6 refer to the number of alternatives available to the subject):

(1) 0, 1, 6, 6, 1, 0
(2) 1, 0, 6, 6, 0, 1
(3) 6, 0, 1, 1, 0, 6
(4) 0, 6, 1, 1, 6, 0
(5) 1, 6, 0, 0, 6, 1
(6) 6, 1, 0, 0, 1, 6.

## *Results*

*Treatment of the Data.*—For each subject, the two time values obtained for each of the conditions (0, 1, and 6 alternatives) were averaged. As in the case of the experiment on perseverance as a function of degree of confidence, the time

[9] It sometimes happened that a solution was achieved with the first card of a given game. When this occurred, the following game was a repetition of the condition.

data for each of the 3 conditions in the present experiment were found to deviate from a normal distribution. For each subject, the raw time-values were transmuted into time-ratio values relative to the time-value for the condition, '6 alternatives,' which was taken as unity. The two distributions of ratio-values thus obtained for conditions '0 alternative' and '1 alternative' were also found to be a-normal.

For purposes of statistical analysis, a logarithmic treatment of the data, as in the case of the preceding experiment, was found most suitable. This treatment yields an excellent correspondence between the empirically determined range of ratio values and the limits as defined by $\pm 3\sigma$, for the condition '1 alternative.' The difference between the upper limits of the 'no alternative' condition as determined empirically (16.50) and statistically (22.38) is amply accounted for by the flatness of the curve and its extremely gradual approach to the baseline. Further data would unquestionably have supplied ratio-values above 16.50.

Table III presents the results for the 42 subjects. The 6 groups differed in the order of presentation of the number of alternatives.

1. There is very little consistency among the subjects as regards the amount of time spent before abandoning an activity which has a stated number of alternatives. The times ranges for each of the three conditions are:

| | |
|---|---|
| No alternative: | 1′45″ – 17′32″ |
| 1 alternative: | 1′ 3″ – 7′30″ |
| 6 alternatives: | 12″ – 5′10″ |

2. A study of Table III reveals a marked tendency among almost all subjects to spend progressively less time before abandoning a first activity, as the number of alternative activities increases from 0 to 1, and from 1 to 6.

The time-ratio ranges, relative to '6 alternatives' taken as 1.00, are:

| | |
|---|---|
| No alternative: | 1.50 – 16.25 |
| 1 alternative: | .78 – 7.83 |

TABLE III

INDIVIDUAL RECORDS OF 42 SUBJECTS:
TIME SPENT BEFORE ABANDONING THE FIRST ATTEMPT AT THE SOLUTION OF PROBLEMS WHEN DIFFERENT NUMBERS OF ALTERNATIVE ATTEMPTS WERE AVAILABLE

Group I

| Subject | Number of Alternatives | | | | | |
|---|---|---|---|---|---|---|
| | 0 | | 1 | | 6 | |
| 1 | 4′ 28″ | 4.12 | 2′ 56″ | 2.71 | 1′ 5″ | 1.00 |
| 2 | 3′ 52″ | 3.09 | 2′ 4″ | 1.65 | 1′ 15″ | 1.00 |
| 3 | 9′ 6″ | 8.40 | 4′ 46″ | 4.40 | 1′ 5″ | 1.00 |
| 4 | 17′ 32″ | 12.67 | 7′ 30″ | 5.42 | 1′ 23″ | 1.00 |
| 5 | 6′ 22″ | 8.13 | 3′ 2″ | 3.87 | 47″ | 1.00 |
| 6 | 12′ 8″ | 5.28 | 3′ 4″ | 1.33 | 2′ 18″ | 1.00 |
| 7 | 4′ 36″ | 3.41 | 2′ 5″ | 1.54 | 1′ 21″ | 1.00 |

Group II

| Subject | Number of Alternatives | | | | | |
|---|---|---|---|---|---|---|
| | 0 | | 6 | | 1 | |
| 8 | 5′ 20″ | 2.83 | 1′ 53″ | 1.00 | 3′ 30″ | 1.86 |
| 9 | 4′ 49″ | 3.61 | 1′ 20″ | 1.00 | 3′ 17″ | 2.46 |
| 10 | 3′ 13″ | 3.38 | 57″ | 1.00 | 1′ 57″ | 2.05 |
| 11 | 7′ 7″ | 7.49 | 57″ | 1.00 | 2′ 38″ | 2.77 |
| 12 | 3′ 10″ | 1.50 | 2′ 7″ | 1.00 | 2′ 15″ | 1.06 |
| 13 | 9′ 47″ | 4.77 | 2′ 3″ | 1.00 | 2′ 9″ | 1.05 |
| 14 | 4′ 35″ | 6.70 | 41″ | 1.00 | 2′ 30″ | 3.66 |

Group III

| Subject | Number of Alternatives | | | | | |
|---|---|---|---|---|---|---|
| | 1 | | 0 | | 6 | |
| 15 | 4′ 6″ | 1.82 | 7′ 20″ | 3.26 | 2′ 15″ | 1.00 |
| 16 | 5′ 50″ | 1.13 | 11′ 57″ | 2.31 | 5′ 10″ | 1.00 |
| 17 | 1′ 34″ | 7.83 | 3′ 15″ | 16.25 | 12″ | 1.00 |
| 18 | 2′ | 1.24 | 9′ 52″ | 6.10 | 1′ 37″ | 1.00 |
| 19 | 1′ 29″ | 1.37 | 7′ 6″ | 6.55 | 1′ 5″ | 1.00 |
| 20 | 2′ 2″ | 2.26 | 2′ 7″ | 2.35 | 54″ | 1.00 |
| 21 | 1′ 52″ | 1.04 | 4′ 27″ | 2.47 | 1′ 48″ | 1.00 |

TABLE III—(*Continued*)

Group IV

| Subject | Number of Alternatives | | | | | |
|---|---|---|---|---|---|---|
| | 1 | | 6 | | 0 | |
| 22 | 4′ 17″ | 1.44 | 2′ 59″ | 1.00 | 8′ 7″ | 2.72 |
| 23 | 3′ 27″ | 1.83 | 1′ 53″ | 1.00 | 5′ 35″ | 2.96 |
| 24 | 7′ 30″ | 2.06 | 3′ 38″ | 1.00 | 15′ 43″ | 4.33 |
| 25 | 1′ 3″ | 1.29 | 49″ | 1.00 | 2′ 20″ | 2.86 |
| 26 | 5′ 13″ | 1.38 | 3′ 47″ | 1.00 | 6′ | 1.59 |
| 27 | 3′ 10″ | 2.18 | 1′ 27″ | 1.00 | 4′ 55″ | 3.39 |
| 28 | 3′ 8″ | 1.79 | 1′ 45″ | 1.00 | 5′ 38″ | 3.22 |

Group V

| Subject | Number of Alternatives | | | | | |
|---|---|---|---|---|---|---|
| | 6 | | 0 | | 1 | |
| 29 | 58″ | 1.00 | 3′ 29″ | 3.60 | 2′ 12″ | 2.28 |
| 30 | 3′ 50″ | 1.00 | 7′ 24″ | 1.93 | 3′ | .78 |
| 31 | 45″ | 1.00 | 4′ 58″ | 6.62 | 2′ 13″ | 2.96 |
| 32 | 38″ | 1.00 | 4′ 15″ | 6.71 | 1′ 51″ | 2.92 |
| 33 | 2′ 15″ | 1.00 | 10′ 18″ | 4.58 | 4′ 13″ | 1.87 |
| 34 | 1′ 5″ | 1.00 | 6′ 8″ | 5.66 | 3′ 47″ | 3.50 |
| 35 | 1′ 18″ | 1.00 | 8′ 30″ | 6.54 | 4′ 48″ | 3.69 |

Group VI

| Subject | Number of Alternatives | | | | | |
|---|---|---|---|---|---|---|
| | 6 | | 1 | | 0 | |
| 36 | 2′ 18″ | 1.00 | 4′ 18″ | 1.87 | 13′ | 5.65 |
| 37 | 2′ 1″ | 1.00 | 3′ 39″ | 1.81 | 4′ 15″ | 2.11 |
| 38 | 1′ 30″ | 1.00 | 3′ 40″ | 2.44 | 7′ 42″ | 5.13 |
| 39 | 1′ 19″ | 1.00 | 3′ 7″ | 2.37 | 11′ 42″ | 8.89 |
| 40 | 1′ 2″ | 1.00 | 4′ 27″ | 4.22 | 9′ 17″ | 8.99 |
| 41 | 45″ | 1.00 | 2′ 3″ | 2.73 | 1′ 45″ | 2.33 |
| 42 | 42″ | 1.00 | 1′ 10″ | 1.67 | 5′ 2″ | 7.19 |

The values to the left of each column represent time, in minutes and seconds.

The values to the right of each column represent time-ratios, relative to condition '6 alternatives' taken as unity.

3. The geometric means of the time-ratios for the two experimental conditions, with the condition '6 alternatives' treated as 100, are:

No alternative: 433.00
1 alternative: 209.06

Table IV gives the differences between the geometric means (in logarithmic terms), the standard deviations of the differences, and the critical ratios. The differences between the means for the three experimental conditions are completely reliable.[10]

TABLE IV

DIFFERENCES BETWEEN THE GEOMETRIC MEANS FOR THE THREE EXPERIMENTAL CONDITIONS, WITH STANDARD DEVIATIONS OF THE DIFFERENCES, AND CRITICAL RATIOS

| Conditions Compared | Difference Between Logs of Geometric Means | S.D.diff. | Critical Ratio |
|---|---|---|---|
| 0 and 1 alternative..... | .316234 | .048 | 6.59 |
| 1 and 6 alternatives..... | .320262 | .032 | 10.01 |

## SUMMARY AND DISCUSSION

While many discussions of human behavior in problem settings have given fragmentary hints about illogical and obstinate stereotypy of response-modes in the face of continued failure, they have shed little light on the factors which make for fixity or fluidity of behavior in such situations. The experiments here reported represent an attempt to isolate two of the factors productive of such behavior. The results of the experiments justify the following conclusions:

1. The amount of time spent in a given activity varies directly with the degree of the individual's assurance that the activity will issue in success. When the chances for the solution of a problem are considered perfect, the time spent on it is strikingly longer than when the chances are all but perfect; similarly, when the chances are viewed as even, the point of resignation is reached still earlier, but definitely later than when the chances are considered very slight. Each difference has been found to be large and consistent. The reliability of the differences between all adjacent pairs of conditions is larger than the statistical criterion for complete reliability. (Experiment I)

[10] The critical ratios are even greater when either the raw times or the time-ratios are treated on the assumption of normal distribution.

2. The amount of time spent in a given activity is an inverse function of the number of alternative and equally promising activities which the individual knows are open to him but holds in abeyance while pursuing the given activity. For the three conditions tested, the results are unequivocal. When the subject has but one means of tackling the problem, his perseverance in it is strikingly greater than it is on the first of two attacks; when he has as many as seven, he abandons the first very easily indeed. Statistical tests show these differences to be thoroughly reliable. (Experiment II)

These conclusions bear an interesting relation to certain observations made by Maier.[11] He presented two groups of subjects with the same problems. He instructed one group to beware of remaining fixed in a single mode of problem-approach, warned them that for a difficult problem the obvious 'direction' was not the correct one, and urged them to be on the lookout for new ways of attacking the problem. The second group was not given this preliminary lecture. Since in the case of each problem, the method of solution was not an 'obvious' one, that is, not one which would be the first adopted by the subject, the attainment of the solution required that the subject abandon the initial mode of attack. Maier found that the number of instructed subjects who solved the problems surpassed that of the control subjects.

This result may be interpreted in the terms of the present findings. Since the solution of Maier's problems was not to be achieved by the pursuit of the course of activity first assumed, anything furthering the termination of that course facilitated solution. Maier's instructions helped to effect the rapid resignation of the course in two ways. In the first place, the instructed subjects were presumably less certain of the success of the course first pursued, since the lecture all but told them to mistrust it. Second, although the subjects were not directly provided with alternative courses of activity, the instructions surely facilitated their occurrence by suggesting their existence and by directing the subjects to seek for

[11] N. R. F. Maier, An aspect of human reasoning, *Brit. J. Psychol.*, 1933, **24**, 144–155.

them. The relationship of these influences to the factors studied in the present investigation is obvious.

The results of Experiment I demonstrate that the degree of an individual's persistence in a given activity varies with the strength of his conviction that the activity will result in satisfaction of his motive. It is suggested that these results may be interpreted in terms of the general principle that the degree of persistence in an activity depends upon the strength of its association with a situation-complex. Such an interpretation may be justified by the fact that in normal experience conviction that an activity will issue in success usually derives, not as in the present experiment from instructions to this effect, but from successful activity in similar situations in the past, that is, from strong association between activity and situation-complex.

In Experiment II, it was found that the subject's knowledge of the availability of alternative modes of attack upon a problem affected the length of time he devoted to the initial attack. These alternative modes of attack may be understood as competing stimulus factors productive of response-sequences different from that produced by the initial attack. If the subject's primary responses, *i.e.*, those determined by the initial attack, prove unavailing, it may be supposed that they will be weakened through failure, and that they will be replaced more readily by other responses, *i.e.*, those determined by the alternative modes of attack, than they would be were no such competing stimulation operative. Therefore, the larger the number of modes of attack which the subject recognizes as possible ways of solving a given problem, the less the chance that his behavior will be long dominated by any single approach.

(Manuscript received May 14, 1940)

# LEVEL OF ASPIRATION AS AFFECTED BY RELATIVE STANDING IN AN EXPERIMENTAL SOCIAL GROUP

BY ERNEST R. HILGARD, EDWARD M. SAIT, AND G. ANN MAGARET

*Stanford University*

The experiments here reported are concerned with the effects of relative success and failure experiences on levels of aspiration in an experimental situation in which the learner is one of several members of a group, all of whose scores are known.[1] Some individuals progressively fall behind the others, some progressively advance ahead of the others. The question to be answered is how these differences in relative standing affect what the individual expects of himself. It might be supposed that repeated success would establish confidence, hence over-prediction, and that failure would produce caution and under-prediction. On the other hand, individuals may refuse to estimate their own performances as much better or worse than the group, and hence be conservative in attributing either success or failure to themselves. Only experimental data secured under a variety of conditions can aid in the interpretation of the circumstances under which one of these results will be found and those circumstances under which the other alternative will be found. In the present study, there is a tendency to judge one's performance somewhat in terms of the group, thus tempering either too much success or too much failure through predicting somewhat nearer the mean.

## Materials, Subjects, and Procedure

Results are reported for brief learning series participated in by undergraduate college students working in groups of from 3 to 6 members. The task chosen was successive subtraction of 3-place numbers. Each student was provided with a pad consisting of 11 sheets of problems, 24 problems on a sheet. While ostensibly similar,

[1] The study was supported in part by a grant to the senior author from the Social Science Research Fund of Stanford University.

the material was of graded difficulty, so that different subjects could be assigned material of different degrees of difficulty without their knowledge. Three sets of problems were used, to be described as easy, medium, and hard. In each set the first three sheets were of equal medium difficulty. The next eight sheets were of progressively easier problems in the easy set, of comparable difficulty throughout for the medium set, and progressively more difficult in the hard set of problems. Difficulty was controlled by the number of problems containing larger digits in the subtrahend than in the minuend, as illustrated below.

| | |
|---|---|
| 398 | 627 |
| − 275 | − 458 |
| Easy | Hard |

That this device sufficed to produce differences in success, particularly as between the easy and hard sets, will be evident from the data to be presented.

In Series A, different members of the same group received different sets of materials, so that relative standing was determined to a considerable extent by the set of materials received, rather than by initial differences in ability.[2] Individual differences in subtraction ability are so wide, however, that rank in the group was not determined entirely by the objective difficulty of the material. In order to provide a control for further analysis, another series, known as Series B, was undertaken, in which all members of a group had materials of the same objective difficulty. In this series relative position was determined by individual differences in ability rather than by objective differences in the problems assigned.

Except for the materials assigned, groups were handled alike in both series. Each group was assembled in a small room with ordinary side-arm classroom chairs arranged facing a table at which the experimenter sat. The subjects were instructed to work as rapidly as possible consistent with accuracy. At a ready signal, all began at once on the first sheet. After the sheet was completed, the subject raised his hand, so that the experimenter could keep a record of the time for that trial, each sheet constituting one trial. When all had finished one sheet, the results for the trial were recorded as follows. The experimenter asked each subject to write down on the back of the sheet an estimate of the time which was taken for the sheet. This score is known as the estimate of past performance (*EPP*).[3] Then the actual time required was announced by the experimenter. This is known as the performance score (*P*). The announcement of the performance score for each member was heard by the other members of the group, scores being read off each time in order from the best to the poorest performance, with a statement of the increase or decrease over the previous trial. Then each subject wrote down on the back of the same sheet his estimate in seconds of the score he thought he would be able to make on the next

[2] It is assumed that the subjects were unaware of differences in their problem-sheets, so that relative success within the group was attributed by them to their own efforts rather than to differences in objective difficulty of materials. There was no evidence to invalidate this assumption. Some of the subjects detected differences in difficulty in successive sheets, but they believed that similar changes occurred on the sheets of the other subjects.

[3] The estimates of past performance have not been used in the analysis of the data to be presented. The potential significance of this measure is discussed in a related study by Hilgard and Sait (3). Various possible meanings of level of aspiration, and various types of measure are considered by Gardner (2).

sheet. This is known as the estimate of future performance (*EFP*). The discrepancy between the estimate of future performance and the previous performance score is the measure of level of aspiration used in this study. The social context under which estimates were made was enhanced (*a*) by having group members work simultaneously, the length of each trial being determined by the slowest member, and (*b*) by having the experimenter read off after each trial the individual performance times. The estimates of past and future performance were made privately, so that there was no public pressure to alter an estimate in order to secure the approval of the group. At the end of the period, the data available consisted of completed problems on 11 sheets from each subject, each sheet containing an estimate of past performance, the actual performance, and an estimate of future performance. While all of the records have been corrected for computational errors, error scores are not reported. Since errors were not known within the time of experimenting, these could have affected estimates of performance only in those cases in which accuracy was deliberately sacrificed in order to increase speed. Errors averaged less than 20 per set of 264 problems, somewhat more errors being made on the hard problems than on the easy ones. Except for two striking cases, the error scores gave evidence that the subjects were attempting to be accurate while working for speed. These two subjects made over 100 errors each, one with hard materials, the other with easy materials. Since these error scores fell outside the distribution of error scores from other subjects, the results from these subjects have been omitted in the calculations.

## Series A. Relative Position in a Group Influenced by Objective Differences in Difficulty of Material

Each of the smaller groups comprising Series A contained members assigned material of different difficulty. Some subjects had to be discarded because they failed to complete the calculations, others through missing estimates on some of the trials. There remain for treatment a total of 18 groups, of which there were 3 subjects in 6 of the groups, 4 subjects in 5 groups, 5 subjects in 6 groups, and 6 subjects in the remaining group. This gives a total of 74 subjects treated.

In order to determine the extent to which position in the group was determined by the difficulty of materials, subjects have been classified as falling within the upper or lower two ranks of the group in which they worked, as based on performance scores during the last four trials, when differences between materials were greatest. This ranking into upper and lower two ranks accounts for four of the subjects in each group of four or more. The remaining subjects are classified together as of medium rank. In the case of groups of three subjects, the highest is placed in the uppermost class, the lowest in the bottom class, and the second member in the

TABLE 1

EXTENT TO WHICH RELATIVE POSITION IN GROUP IS DETERMINED BY DIFFICULTY OF MATERIAL

| | Number of Subjects Falling in each Position in Group, Trials 8–11* | | | |
|---|---|---|---|---|
| | Ranks 1 and 2 | Middle Ranks | Ranks Last and Next-to-Last | Total |
| Easy material | 18 | 4 | 2 | 24 |
| Medium material | 7 | 5 | 13 | 25 |
| Hard material | 5 | 5 | 15 | 25 |
| | | | | 74 |

* These are the trials in which differences in difficulty are at a maximum.

middle rank. The results are shown in Table 1. It is evident from the table that those with easy material were predominantly in the upper ranks, those with hard material predominantly in the lower ranks. By this classification, medium materials resulted in rankings nearer to hard than to easy materials.

Quantitative results for the actual and estimated scores are presented in Table 2, and the data for the extreme groups

TABLE 2

LEVEL OF ASPIRATION AS REPRESENTED BY DISCREPANCY BETWEEN ACTUAL PAST PERFORMANCE AND ESTIMATE OF FUTURE PERFORMANCE DURING DIFFERENT STAGES OF PRACTICE (SERIES A)

Past performance, *P*; estimate of future performance, *EFP*

| | Trials | No. of Subjects | Mean Scores in Seconds | | | | |
|---|---|---|---|---|---|---|---|
| | | | *P* | *EFP* | *P–EFP* | $\sigma_D$ | $D/\sigma_D$ |
| *E*-subjects, with easy materials | 1– 3 | 24 | 111.6 | 105.1 | + 6.5 | 2.2 | + 3.0* |
| | 4– 7 | | 91.2 | 92.3 | – 1.1 | 1.0 | – 1.1 |
| | 8–11 | | 79.3 | 82.6 | – 3.4 | 1.0 | – 3.4 |
| *M*-subjects, with materials of medium difficulty | 1– 3 | 25 | 112.4 | 108.6 | + 3.8 | 1.8 | + 2.1 |
| | 4– 7 | | 102.7 | 101.2 | + 1.5 | 1.3 | + 1.1 |
| | 8–11 | | 100.2 | 99.2 | + 1.0 | 1.2 | + 0.8 |
| *H*-subjects, with hard materials | 1– 3 | 25 | 111.6 | 107.8 | + 3.8 | 1.6 | + 2.4 |
| | 4– 7 | | 109.0 | 107.1 | + 1.9 | 1.2 | + 1.6 |
| | 8–11 | | 108.5 | 104.3 | + 4.2 | 1.6 | + 2.6 |

* A plus (+) value means that the score for the future performance is estimated to be superior to that of the past performance, *i.e.*, the future performance is expected to require less time.

plotted in Fig. 1. In each case the discrepancy between the actual performance and the estimate of future performance is calculated. This discrepancy score is a measure of the stated momentary level of aspiration. If the discrepancy is positive, it means that improvement is anticipated; if the discrepancy score is negative, a loss is expected. Scores are reported in seconds, so that a better score is numerically smaller, and a poorer score numerically larger.

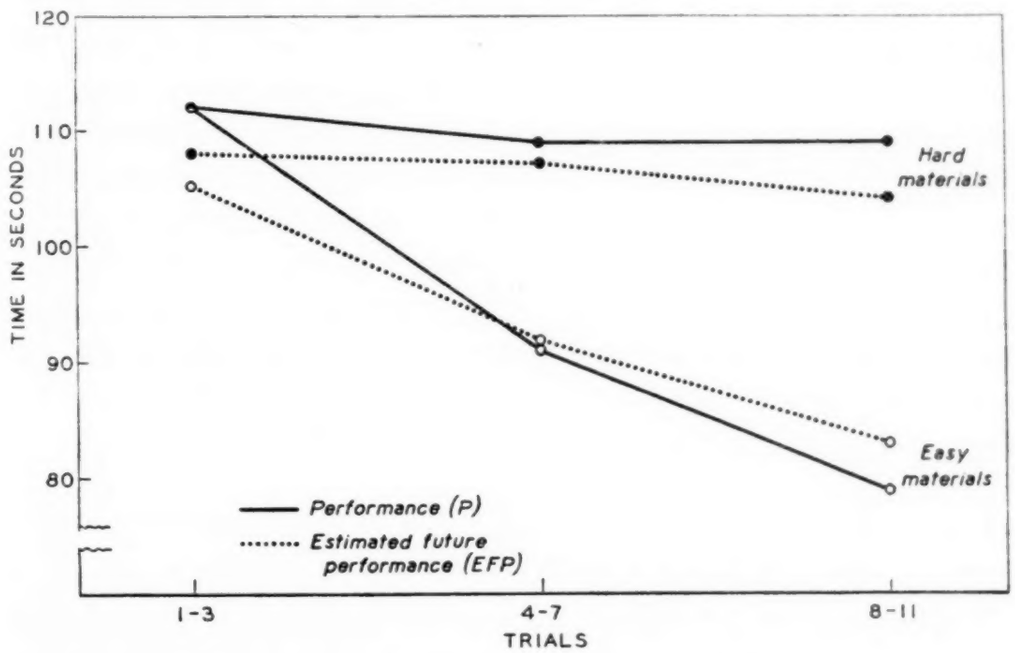


FIG. 1. Performance and estimated future performance for subjects of equal initial ability with material of progressively different objective difficulty. Materials of equal difficulty during trials 1–3, progressively harder and easier, respectively, in later trials. Mean values from 24 subjects with easy materials, 25 subjects with hard materials, data from Table 2. Note the tendency for those with hard material to continue to expect to improve, while late in practice those with easy material expect to do less well (*i.e.*, to require more time).

The main trends can best be observed in Fig. 1, in which results are plotted for *E*-subjects (with easy material) and for *H*-subjects (with hard material). There is a consistent tendency throughout for those with hard materials to expect their scores to improve, so that the estimated performance (in terms of time to be required) falls below the past performance. On the other hand, those with easy materials, at first expecting to improve, end by expecting to lose ground on the next trial, so that their estimates of future performance at

the end lie above the actual performance. In other words, so far as group averages are concerned, the unsuccessful subjects refuse to concede defeat, and the successful ones refuse to accept fully their superiority. The extent to which the mean results represent the groups as a whole is indicated by the critical ratios of Table 2. It is evident that there are wide individual differences, but the subjects are too few to make further fractionation of results useful.

Trials 1–3 are in the nature of controls on performance, since they provide scores for all subjects with materials of comparable difficulty. The plan of the experiment requires that the subjects should be of comparable ability initially, so that differences in performance (and hence in aspiration level) may be created artificially in the experimental situation. The three groups are roughly comparable in Trials 1–3, so that this demand of the experiment appears to be met. Differences introduced by changes in difficulty succeeded in producing group differences in expressed levels of aspiration, even though the groups were essentially alike at the start.

### Series B. Relative Position in a Group Independent of Objective Difficulty of Material

Since college students come to an arithmetic experiment with estimates of their own success in arithmetical operations, it is conceivable that persistent individual differences may be more important in determining both scores and aspiration levels than the objective difficulties of the problems presented. In Series A it was shown that group differences could be produced through relative difficulty of materials. Series B was designed as a control, to see to what extent these differences were a function of the social situation. If the social situation, described objectively by the relative position in the group, is the important determiner, then results corresponding to those of Series A should be found for subjects of high, medium, and low ranks relative to their fellows, even though the material is of uniform difficulty. Data are treated from 11 groups under these conditions, 1 with 4 members, 5 with 5 members, and 5 with 6 members, a total of 59 subjects.

In order to secure scores falling within the same range as those of Series A, groups were practiced with uniform material throughout for any one group, but some with hard, others with medium, and still others with easy material. Subjects were then classified into three degrees of success in terms of relative position in the group within the final four trials, corresponding to the classification of the subjects of Series A in Table 1. This resulted in 22 of the 59 subjects falling in high ranks, 22 in the low ranks, and 15 in the middle ranks. The ranking is determined in Series B by the individual abilities of the subjects, since material averaged of equal objective difficulty for each rank. Results for subjects combined according to the three relative ranks are presented in Table 3, and the results for the extreme groups plotted in Fig. 2.

TABLE 3

LEVEL OF ASPIRATION AS REPRESENTED BY DISCREPANCY BETWEEN ACTUAL PAST PERFORMANCE AND ESTIMATE OF FUTURE PERFORMANCE DURING DIFFERENT STAGES OF PRACTICE (SERIES B)

Past performance, *P*; estimate of future performance, *EFP*

| | Trials | No. of Subjects | Mean Scores in Seconds | | | | |
|---|---|---|---|---|---|---|---|
| | | | *P* | *EFP* | *P–EFP* | $\sigma_D$ | $D/\sigma_D$ |
| *A*-subjects, falling in ranks 1 and 2 within trials 8–11 | 1– 3 | 22 | 96.6 | 95.2 | + 1.4 | 1.1 | + 1.3* |
| | 4– 7 | | 84.4 | 86.9 | − 2.5 | 0.8 | − 3.2 |
| | 8–11 | | 80.2 | 81.4 | − 1.2 | 1.1 | − 1.1 |
| *B*-subjects, falling in middle ranks within trials 8–11 | 1– 3 | 15 | 111.4 | 107.0 | + 4.3 | 1.0 | + 4.0 |
| | 4– 7 | | 97.9 | 98.6 | − 0.7 | 1.0 | − 0.7 |
| | 8–11 | | 96.0 | 94.5 | + 1.4 | 0.6 | + 2.2 |
| *C*-subjects, falling in last two ranks within trials 8–11 | 1– 3 | 22 | 133.9 | 127.7 | + 6.2 | 0.8 | + 7.8 |
| | 4– 7 | | 117.0 | 115.8 | + 1.3 | 1.3 | + 1.0 |
| | 8–11 | | 111.7 | 109.0 | + 2.6 | 1.2 | + 2.3 |

* A plus (+) value means that the score for the future performance is estimated to be superior to that of the past performance, *i.e.*, the future performance is expected to require less time.

Like the subjects of Series A, the successful subjects of Series B interpret their successes cautiously, and estimate that the next scores will be less successful, while the unsuccessful subjects continue to estimate that their scores will improve. In this series, the differences in performance are present ini-

tially, and nothing fictitious has been introduced. The same method of announcing scores was followed, however, so that each subject has knowledge of the scores being made by the others prior to his own estimates. Since results are in general agreement for the two series, it may be conjectured that

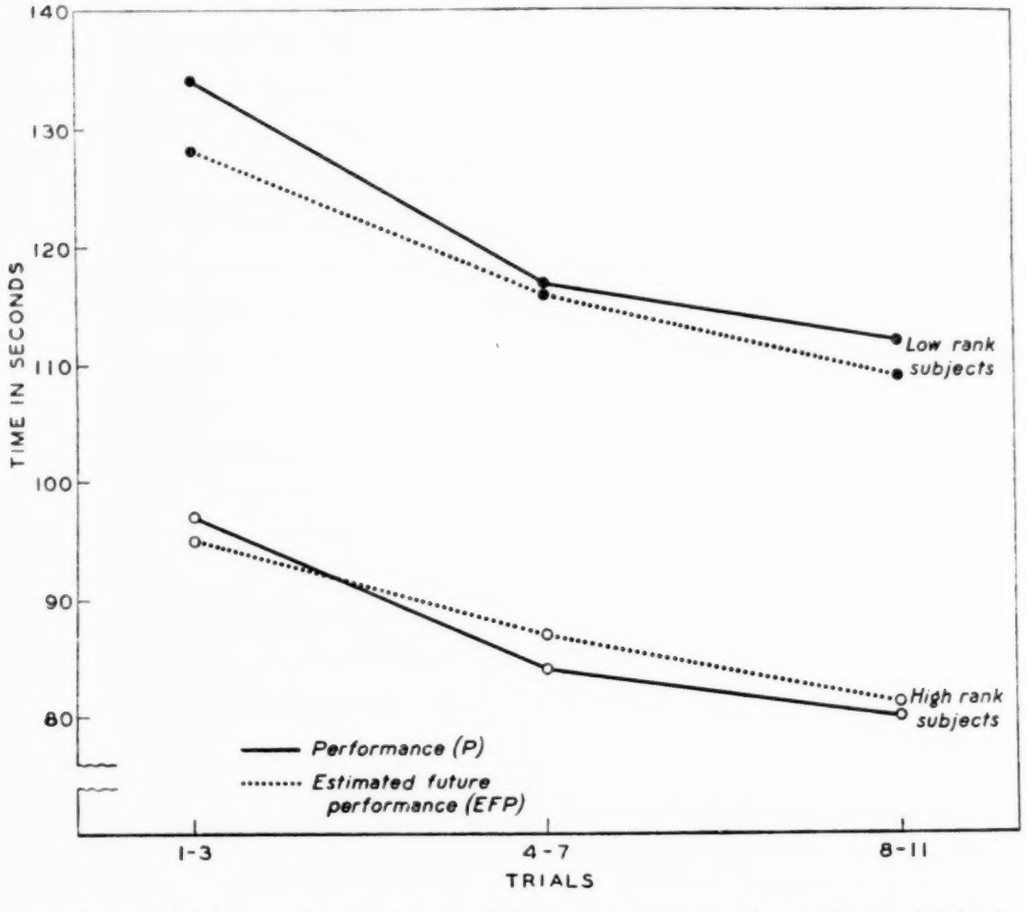


FIG. 2. Performance and estimated future performance for subjects of differing ability with material of equal objective difficulty. The subjects are classified on the basis of ranking in performance in their experimental groups. Mean values from 22 subjects in each group, data from Table 3. Note the tendency for those of low rank to continue to expect to improve, while late in practice those with high rank expect to do less well (*i.e.*, to take longer).

relative position in the group is important in determining the direction of deviation of the estimate from a purely realistic prediction.

The subjects involved are too few and the measures too restricted to permit a fruitful analysis of the data for individual differences. It is evident from the small critical ratios in Table 3 that not all of the successful and unsuccessful subjects

showed the mean trends of the group to which they belonged. It was not the purpose of the present experiments to study level of aspiration as a personality characteristic, however, and no further information was gathered with respect to the individuality of the subjects.

Departures from realistic predictions of future scores are influenced by many factors. Success and failure experiences depend in part upon the 'ego-involvement' in the task, and no ratings of this were made in the experiments. Estimates of future performance may be based upon the expected trend of scores and upon the previous *best* score, as well as upon the immediately preceding score from which the discrepancy is calculated. In the case of those with material getting progressively harder, estimates based on best scores would produce the effect found, namely, a prediction of scores better than those being made. On the other hand, prediction in terms of best score cannot explain why those with easy material predict *poorer* scores for themselves. In view of the many factors which might distort an individual's reactions in this situation, it is rather remarkable that trends are as consistent as they are.

## An Attempted Interpretation of the Tendency for Estimates of Future Performance to Regress toward the Mean of the Social Group

In the two series of experiments reported, subjects ranking superior in relation to their social group tended to estimate their future performances too low, while those subjects making inferior scores tended to estimate their future performances too high. Although critical ratios were so low that caution is needed in interpretation, the trends within the groups were clear. A similar observation has been reported by Anderson and Brandt (1) in level of aspiration experiments with school children, but in a situation requiring public announcement of anticipated scores.

When the estimates of future performance are made publicly, the desire for social conformity might well produce some regression of predicted scores toward the mean. One

with a good score does not wish to brag, and so asserts that it was a matter of good luck and next time the score will be lower. Similarly, one with a low score does not wish to admit that he is a deviate, attributes the low score to bad luck, and expresses an optimistic estimate. It must be pointed out that the discrepancies are in fact of small size, so that the successful subject still claims a superior performance, and the unsuccessful subject still predicts a poor performance. But the direction of departure of estimates from actual scores is consistent enough to be psychologically interesting.

When the estimates are not made publicly, the question arises whether the same tendency may not be at work. That is, a subject may still think of the social approval of the group when writing his score privately, just as a person may feel some embarrassment at voting for himself on a secret ballot. One does not depart completely from the cultural environment when he is alone. In these experiments, the record was to become the property of the experimenter, so that some indirect social pressure might have been present; more, that is, than if the material were strictly anonymous. Further experimentation is necessary to determine the differences to be expected in stated estimates of future performance with different degrees of publicity accompanying the estimates.

## Summary

1. Experiments were conducted with college students working in small groups consisting of 3 to 6 members. The task assigned was successive subtraction. After each trial the experimenter announced the time required, and the subject made a private estimate of the time which would be required on the next trial. The measure of level of aspiration used was the discrepancy between the performance just completed and the estimated future performance.

2. In Series A, materials were of equal medium difficulty in the first 3 trials, but there followed 8 trials in which for a third of the subjects the materials became easier; for a third, more difficult; while for the last third they remained the same. Subjects were unaware of the differential difficulty introduced.

Results with 74 subjects showed a tendency for those with easy materials to estimate their performances poorer than they actually were, and for those with hard materials to estimate their performances better than they actually were. A tendency is thus shown for estimates to regress slightly toward the mean performance of the social group.

3. In Series B, materials were all of equal difficulty for the members of a group, so that relative position was determined by arithmetical ability rather than by introduced differences in material. In experiments with 59 subjects, trends similar to those of Series A were shown.

4. Since estimates were made privately, group pressure can be used to explain the tendency to estimate towards the mean of the group only if the desire for social conformity is somehow internalized.

(Manuscript received May 20, 1940)

## References

1. Anderson, H. H., and Brandt, H. F., Study of motivation involving self-announced goals of fifth grade children and the concept of level of aspiration, *J. soc. Psychol.*, 1939, **10**, 209–232.
2. Gardner, J. W., The use of the term 'level of aspiration,' *Psychol. Rev.*, 1940, **47**, 59–68.
3. Hilgard, E. R., and Sait, E. M., Interrelations of measures of level of aspiration in two perceptual-motor tasks (In preparation).

# AN EXPERIMENTAL INVESTIGATION OF CHANGES IN THE MEANING OF LEVEL OF ASPIRATION

BY ROSALIND GOULD AND HELEN BLOCK LEWIS

*Brooklyn College*

In the main, studies on aspiration level have been concerned with the nature of individual differences in internal standards of achievement as they revealed themselves in the stated estimates of future performance. In such experiments, the subject is unaware of his relative standing in performance although he is well aware that others are participating in the study and that some 'average' is being established. Most subjects voice their concern with this unknown 'average,' but the operation of such an external standard is necessarily a diffuse one, acting as an added anxiety spur to greater effort, rather than as a specific influence on the magnitude of the aspiration estimates themselves. Apart from the problem of individual differences in internal standards, it would be important to determine the extent and kind of influence exerted by specific external standards on aspirations. The effect of any given external standard would, of course, be a function not only of the nature of the standard but of a deeper-lying self-evaluation which leads the individual to caution or self-confidence in many fields of work.

Some of the recent experiments on aspiration level have introduced specific external standards. Gardner (**3**) devised a technique whereby the same performance scores in percentile figures are announced to all the subjects, thus imposing on all the same objective external standard. Such a procedure does not, of course, assure all subjects' imputing the same meaning to these standards of proficiency, just as two subjects obtaining the same performance score in the more unstructured aspiration-level situation, do not necessarily regard their achievements in the same way. Gardner was not interested, however, in the problem of the effect of external standards on

aspiration estimates, but rather with the creation of an identical objective situation for all subjects. Anderson and Brandt (1) introduced an external standard by informing each student of his relative achievement-status in comparison with his classmates. This knowledge produced differences in the size of the difference-score directly in accordance with the performance status.

In these studies, the problem of individual differences in response to such standards has been largely neglected. Hertzman and Festinger (5) and Gould (4) have set up experimental situations where the external standard introduced has been especially adapted to the nature of the individual's aspiration-level as revealed in an initial or preliminary session of the usual aspiration level experiment. A direct attack on the study of the influence of different external standards upon aspiration estimates was recently made by Chapman and Volkmann (2). They, too, were not concerned with individual differences, but with differences between groups exposed to different standards of achievement. Their technique involved a comparison of the influence of external standards of reference without knowledge of performance, with the influence of external standards accompanied by knowledge of performance. Their results indicated an 'anchoring' effect of self-knowledge such that the external standard was relatively impotent in the face of it. Unfortunately, this study was solely a quantitative one leaving to speculation the bases of the obtained differences. The exact way in which the external standard operated, or failed to operate was not clear. Does knowledge of another group's performance actually change the subject's own goal? Does the presented standard of reference merely increase or decrease the anxiety of the subject without changing the nature of the goals set up under the different conditions? Or perhaps the external standard simply alters the meaning of the expressed aspiration, rather than its size, changing it from an estimate to a distant hope or from a nebulous possibility to a necessary achievement. In order to answer these questions, the present experiment was performed.

## Procedure

Our procedure attempted a combination of group experimentation with individual analysis. Three different groups of Brooklyn College students in Introductory Psychology were studied under three different conditions. (Our assumption was, of course, that the groups were comparable.) All of our subjects were given from 6 to 8 trials of 'General Information' tests which had previously been equated for difficulty. This procedure differs from Chapman and Volkmann's who used but four trials, only two of which were subject to the influence of an external standard. Each test contained 20 items and the subjects were allowed 3 minutes in which to complete their answers. Previous work with the test had shown that 3 minutes was ample time for filling in what answers the subject knew.

After the initial trial, each subject in the group received a prepared sheet containing his name, his actual score on the last performance trial, an indication of the relative standing of that score, and the question to be filled in *before* the test began, "What score do you intend to make next time?" Group A containing 66 individuals served as *Control.* The relative standing of each of these subjects was presented in terms of a 5 point scale: very much above average, above average, average, below, or very much below the average of his own group of classmates. Group B, containing 63 individuals, served as the first experimental group. The relative standing of each subject here was phrased as above average, average, or below the average of 'a group of college professors' whose average performance score was announced as 12. Group C, containing 51 subjects, served as the second experimental group. Each subject, here, had his relative standing revealed as in Group B, but this time phrased as above average or below the average of 'a group of unselected WPA workers,' whose average score was also announced as 12.[1] This average of 12 was approximately 2 points above that of the actual average of the subjects themselves. There were approximately an equal number of men and women in Groups A, B, and C. Shortly after the close of the experiment, about 70 selected at random from our 180 subjects were interviewed on the following questions: 1. What did you think was the purpose of this experiment? 2. How did you arrive at your estimate? 3. What was the basis of your estimate? 4. How did you feel when your estimate and performance did not agree? 5. Did you have a feeling of failure? Why? Did you have a feeling of success? Why? 6. How do you react to failure in general, *e.g.*, final exams, etc.? 7. What do you think of WPA workers as a group? 8. What did you think of the average of 12 made by WPA workers (college professors)? Did this average influence your estimate? 9. Had you known that the score of 12 was obtained by WPA workers (college professors) do you think your estimate would have been changed?

## Results and Discussion

### 1. *Trial-by-trial Analysis*

A trial-by-trial account of the average performances, average estimates and average discrepancy score[2] obtained

[1] Henceforward the control condition will be referred to by the symbol, 'Control,' the 1st experimental condition by the symbol 'Profs' and the 2nd experimental condition by the symbol 'WPA.'

[2] Discrepancy score indicates the difference between performance score and aspiration.

by men and women separately and together under the three experimental conditions is presented in Table 1. The odd-even reliability coefficients for aspiration and performance both are high, .79 and .81, respectively.[3] Table I shows that,

TABLE I

TRIAL-BY-TRIAL ANALYSIS OF PERFORMANCES, ASPIRATION, AND D-SCORE UNDER 3 CONDITIONS FOR MEN AND WOMEN SEPARATELY AND FOR BOTH COMBINED

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | N |
|---|---|---|---|---|---|---|---|---|---|
| Performance | | | | | | | | | |
| Control | | | | | | | | | |
| M..... | 8.13 | 9.77 | 11.50 | 10.45 | 10.90 | 9.88 | 9.58 | 9.86 | 26–31 |
| W..... | 6.71 | 7.19 | 7.62 | 7.66 | 9.26 | 8.13 | 6.34 | 7.14 | 26–35 |
| Total.... | 7.42 | 8.44 | 9.50 | 9.03 | 10.03 | 9.03 | 7.79 | 8.50 | 52–66 |
| Prof | | | | | | | | | |
| M..... | 8.27 | 11.14 | 11.70 | 10.93 | 12.55 | 11.17 | 11.50 | 10.78 | 27–31 |
| W..... | 7.87 | 7.78 | 9.19 | 9.21 | 9.90 | 8.48 | 8.31 | 8.24 | 26–32 |
| Total.... | 8.07 | 9.46 | 10.42 | 10.02 | 11.22 | 9.78 | 9.80 | 9.46 | 53–63 |
| WPA | | | | | | | | | |
| M..... | 8.33 | 10.71 | 11.17 | 11.71 | 11.89 | 10.17 | 9.52 | 10.00 | 24–29 |
| W..... | 8.53 | 8.24 | 9.77 | 10.33 | 10.70 | 8.83 | 8.07 | — | 27–30 |
| Total.... | 8.44 | 9.38 | 10.45 | 11.04 | 11.32 | 9.84 | 10.52 | | 51–59 |
| Aspiration | | | | | | | | | |
| Control | | | | | | | | | |
| M..... | 12.17 | 12.00 | 13.07 | 12.39 | 12.97 | 12.11 | 12.96 | | |
| W..... | 8.88 | 8.37 | 8.59 | 8.80 | 10.33 | 9.69 | 8.65 | — | |
| Total.... | 10.64 | 10.18 | 10.72 | 10.48 | 11.67 | 10.77 | 10.73 | | |
| Prof | | | | | | | | | |
| M..... | 12.44 | 13.03 | 13.03 | 13.00 | 13.87 | 13.50 | 13.32 | | |
| W..... | 11.35 | 10.90 | 11.55 | 11.19 | 11.90 | 11.62 | 11.62 | — | |
| Total.... | 11.91 | 11.93 | 12.21 | 12.10 | 12.85 | 12.50 | 12.46 | | |
| WPA | | | | | | | | | |
| M..... | 14.12 | 14.21 | 14.44 | 14.79 | 14.00 | 14.62 | 14.18 | | |
| W..... | 12.07 | 11.07 | 11.92 | 11.85 | 12.96 | 12.11 | — | — | |
| Total.... | 13.27 | 12.59 | 13.18 | 13.37 | 13.48 | 13.39 | | | |
| D-Score | | | | | | | | | |
| Control | | | | | | | | | |
| M..... | 4.04 | 2.23 | 1.57 | 1.94 | 2.07 | 2.23 | 3.38 | | |
| W..... | 2.17 | 1.18 | 0.97 | 1.14 | 1.07 | 1.56 | 2.31 | — | |
| Total.... | 3.22 | 1.74 | 1.22 | 1.45 | 1.64 | 1.74 | 2.94 | | |
| Prof | | | | | | | | | |
| M..... | 4.17 | 1.89 | 1.33 | 2.07 | 1.22 | 2.39 | 1.82 | | |
| W..... | 3.48 | 3.12 | 2.36 | 1.98 | 2.00 | 3.14 | 3.31 | — | |
| Total.... | 3.84 | 2.47 | 1.85 | 2.08 | 1.63 | 2.72 | 2.66 | | |
| WPA | | | | | | | | | |
| M..... | 5.79 | 3.50 | 3.27 | 3.08 | 2.11 | 4.45 | 4.63 | | |
| W..... | 3.54 | 2.83 | 2.15 | 1.52 | 2.26 | 3.78 | — | — | |
| Total.... | 4.83 | 3.21 | 2.73 | 2.33 | 2.16 | 3.55 | — | | |

[3] These are uncorrected coefficients. The size of the correlations is limited by the small spread of the distributions.

within each condition, average performance varies as much as from 3 to 4.5 points, while the variation of average aspiration is only 0.79 in WPA group, 1.33 in Profs group, and 1.07 in the Control group. The maximum difference in performance between control and experimental groups is always smaller than the maximum difference between two trials in any one condition. In other words, the variations of performance within the 8 trials are larger than the differences in ability between our three groups. This is in striking contrast to the differences among the three groups in aspirations in comparable trials. Here, differences are found which are 3 times the size of the variation within each condition. The trend in the size of the estimates is, with the exception of but two trials, clearly in the direction of highest estimates for WPA and lowest for Control, although the size of these differences is not always large. The aspiration estimates, extremely stable from trial-to-trial but varying in magnitude with our 3 conditions, thus clearly reveal the influence of the externally imposed standards. The discrepancy-scores show this influences less clearly, since the variations in performance from trial to trial acted to obscure the fairly consistent level of estimates. As might be expected, for each group, the largest D-scores [4] are obtained in the first trial. One also notes that, with the exception of the WPA group, there is a tendency for the D-scores to be smaller in the later trials, than in the earlier trials.

### 2. *Sex Differences*

An interesting sex difference in D-scores is also apparent from Table I. Under WPA and Control conditions, the D-score for women is consistently lower than for men, whereas in the Profs group, the women show higher D-scores in 5 of the 7 trials. It is difficult to attribute these differences conclusively to a sex difference, since there are differences between the average performance of men and women under the three conditions. On the other hand, the women are consistently inferior to men in performance, so that the patterns

[4] In the future, the symbol 'D-score' will designate the average discrepancy score.

of D-scores and performance differences are not consistently related to each other.

Table II shows the mean performance, aspiration and D-score under three conditions for all 6 trials, for men,

TABLE II

MEAN SCORES AND σ'S FOR MEN, WOMEN AND TOTAL GROUPS IN PERFORMANCE, ASPIRATION AND D-SCORE UNDER 3 CONDITIONS: CONTROL, 'PROF' 'WPA'

| Condition | | Mean Score | σ Dist. |
|---|---|---|---|
| Performance | (N) | | |
| Control Men | (31) | 10.03 | 2.79 |
| Control Women | (35) | 7.53 | 1.90 |
| Total | (66) | 8.74 | 2.67 |
| 'Prof' Man | (31) | 11.01 | 2.70 |
| 'Prof' Women | (32) | 8.64 | 2.57 |
| Total | (63) | 9.80 | 2.89 |
| 'WPA' Men | (28) | 10.45 | 2.43 |
| 'WPA' Women | (23) | 9.17 | 2.62 |
| Total | (51) | 10.12 | 2.61 |
| Aspiration | | | |
| Control Men | | 12.52 | 1.17 |
| Control Women | | 9.05 | 1.83 |
| Total | | 10.74 | 2.47 |
| 'Prof' Men | | 13.17 | 2.09 |
| 'Prof' Women | | 11.45 | 1.77 |
| Total | | 12.29 | 2.13 |
| 'WPA' Men | | 14.35 | 1.76 |
| 'WPA' Women | | 11.99 | 3.02 |
| Total | | 13.25 | 2.68 |
| Difference Score | | | |
| Control Men | | 2.49 | 2.25 |
| Control Women | | 1.52 | 1.18 |
| Total | | 2.00 | 1.87 |
| 'Prof' Men | | 2.16 | 1.40 |
| 'Prof' Women | | 2.81 | 2.04 |
| Total | | 2.49 | 1.83 |
| 'WPA' Men | | 3.90 | 1.75 |
| 'WPA' Women | | 2.82 | 2.15 |
| Total | | 3.13 | 2.08 |

women and for both sexes combined. Table III shows the differences between the means and the reliabilities of these differences. Under all three conditions, the performance scores of men are higher than for women, as are the aspirations. Two of the three performance differences in mean score between men and women are statistically reliable; all of the 3 differences in mean aspiration scores are reliable. Differences between D-scores of men and women are not statistically reliable. Under Control and WPA conditions,

TABLE III

DIFFERENCES BETWEEN MEAN PERFORMANCE, ASPIRATION AND D-SCORE, UNDER 3 CONDITIONS: CONTROL, PROFS, WPA, FOR TOTAL GROUP, MEN, WOMEN AND MEN-WOMEN

| | | Difference Between Means | $D/\sigma D$ |
|---|---|---|---|
| Total Group | Performance | | |
| | Control-Prof. . . . . | 1.06 | 2.16 |
| | Control-WPA. . . . . | 1.38 | 2.88 |
| | WPA-Prof. . . . . | 0.32 | 0.62 |
| | Aspiration | | |
| | Control-Prof. . . . . | 1.55 | 3.78 |
| | Control-WPA. . . . . | 2.51 | 5.23 |
| | WPA-Prof. . . . . | 0.96 | 2.09 |
| | Diff.-Score | | |
| | Control-Prof. . . . . | 0.49 | 1.36 |
| | Control-WPA. . . . . | 1.13 | 3.14 |
| | WPA-Prof. . . . . | 0.64 | 1.73 |
| Men | Performance | | |
| | Control-Prof. . . . . | 0.98 | 1.40 |
| | Control-WPA. . . . . | 0.42 | 0.62 |
| | WPA-Prof. . . . . | 0.56 | 0.84 |
| | Aspiration | | |
| | Control-Prof. . . . . | 0.65 | 1.51 |
| | Control-WPA. . . . . | 1.83 | 4.69 |
| | WPA-Prof. . . . . | 1.18 | 2.36 |
| | D-Score | | |
| | Control-Prof. . . . . | 0.33 | 0.69 |
| | Control-WPA. . . . . | 1.41 | 2.39 |
| | WPA-Prof. . . . . | 1.74 | 4.14 |
| Women | Performance | | |
| | Control-Prof. . . . . | 1.11 | 1.98 |
| | Control-WPA. . . . . | 1.64 | 2.60 |
| | WPA-Prof. . . . . | 0.53 | 0.75 |
| | Aspiration | | |
| | Control-Prof. . . . . | 2.40 | 5.45 |
| | Control-WPA. . . . . | 2.94 | 4.20 |
| | WPA-Prof. . . . . | 0.54 | 0.77 |
| | D-Score | | |
| | Control-Prof. . . . . | 1.29 | 3.15 |
| | Control-WPA. . . . . | 1.30 | 2.65 |
| | WPA-Prof. . . . . | 0.01 | .00 |
| Men-Women | Performance | | |
| | Control. . . . . | 2.50 | 4.24 |
| | Prof. . . . . | 2.37 | 3.54 |
| | WPA. . . . . | 1.28 | 1.80 |
| | Aspiration | | |
| | Control. . . . . | 3.47 | 9.63 |
| | Prof. . . . . | 1.72 | 3.51 |
| | WPA. . . . . | 2.36 | 3.32 |
| | D-Score | | |
| | Control. . . . . | 0.97 | 2.15 |
| | Prof. . . . . | 0.65 | 1.59 |
| | WPA. . . . . | 1.08 | 1.93 |

women have a lower D-score, under the Profs condition, women have a higher D-score.

An interesting sex difference in the relative 'strength' of our two external standards appears also in these tables. It will be noted that the D-score for women in the Profs group is reliably higher than the D-score of women in the Control group. In the men's group, on the other hand, the largest D-score is to be found under the WPA condition. This D-score is higher than that of the men in the Control group (not reliably) and higher than the D-score of men on the Profs group (reliably). Apparently, the WPA standard exerts more of an influence on the men's aspiration level, while the Profs standard is more potent for the women. These differences gain increased importance when they are contrasted with differences in performance between men and women. Women are poorer in performance than men under the Profs standard, but their performance is also poorer than the men's in the WPA condition. Here again we see that differences in D-score cannot be related to consistent differences in performance between men and women.

The interview offers some suggestions to account for these sex differences in response to WPA and Profs standards. Many women, certainly more women than men, are preparing for and vitally concerned with future positions associated with the teaching profession. They did not anticipate WPA as a vocational eventuality as much as did the men. Many men reacted to the WPA in a more personal light spontaneously mentioning friends who were employed on WPA, sarcastically commenting on the 'luck' in being able to get such employment, and apparently greatly concerned with their own ability to escape such 'luck.' Though not many of them considered WPA workers as inferior, it was, in the light of their anxieties concerning future jobs, more pertinent to them to demonstrate ability above that displayed by the 'average of an unselected WPA group.' In a more general vein, one can point out that girls tend more to identify themselves with their teachers, are more academically docile than boys. Thus, in terms of the way they viewed their future

vocational possibilities, a greater competitiveness was aroused by the WPA for men and by the Profs for women.

*Summary*

1. Except in the Profs group, women show lower D-scores. The significance of these differences is obscured somewhat by differences in performance between men and women.

2. None of the differences in D-scores between men and women are statistically reliable.

3. In the Profs group, the D-score for women is reliably higher than the D-score for women under Control conditions.

4. For the men, the D-score in the WPA group is higher than in Control, and reliably higher than Profs.

5. An explanation for these differences is offered from the interview material.

### 3. *Differences among the Three Conditions: Men and Women Combined*

It will be seen from Tables II and III that there are no statistically reliable differences in performance among the groups operating under the three conditions. It is interesting, however, that the two groups operating under an external standard showed better performance scores than the control group, and that the best total average performance of the three groups was that obtained under WPA conditions. The aspiration scores obtained under Profs and WPA conditions are reliably differentiated from that of the Control scores, but not from each other. The WPA mean aspiration is higher than that of the Profs ($D/\sigma D = 2.09$), so that a hierarchy of mean aspirations is formed, the lowest being that of the Control, the highest, WPA. This hierarchy follows the performance hierarchy exactly, but in all cases the aspiration differences are considerably larger than the performance differences.

From these figures alone, it is impossible to determine whether the increase in mean aspiration estimate is a simple function of higher performance level, or whether both hierarchies are functions of the same basic condition of increased incentive. On the surface, the latter possibility seems the

less probable, since our general information test is not responsive to the effects of practice or increased effort. From our interviews with our subjects, however, we have some suggestion that our 8 information tests formed a fairly important part of our subjects' lives during the 4 week period of administration, and that this personal involvement was particularly keen for those groups who were 'competing' with the Profs and WPA workers. If this was the case, then it is possible that the 'bubbling-over' with information which we observed during our experiment operated in our two experimental conditions to make both aspiration and performance higher than under control conditions.

The total mean D-scores follow the same trend, 2.00, 2.49 and 3.13, but the only reliable difference is between Control and WPA groups. Apparently the disparity between performance and aspiration increases from the Control to WPA conditions and is greater with both external standards than for the Control groups.

*Summary*

1. No statistically reliable differences in performance were found for the total groups among our three conditions, though the differences in performance between Control and WPA groups approaches reliability.

2. Aspiration scores, following the same order of magnitude, show larger differences (reliable) among the three groups.

3. The D-score is lowest for Control and highest for WPA.

### 4. *Performance, Aspiration, and D-score by Performance Groups*

Table IV shows the mean performance, aspiration and D-score under the three conditions when our subjects are divided into four performance categories: I, those making scores from 5–8; II, 8–10; III, 10–12; and IV, 12 and above. Under all conditions, differences in mean performance of the 4 performance categories are statistically reliable. In general, mean D-scores are lower in the better or higher performance categories, although not all differences in aspiration between performance groups are reliable. The difference in aspiration

between categories I and IV is reliable for all three conditions.[5] For the Control group, the difference in aspiration between the II (8–10) and III (10–12) categories is reliable; for the Profs the difference in aspiration between III (10–12) and IV (12 and above) is reliable. This is also true for the WPA group.

It seems to us most significant that statistically reliable differences in aspiration occur at just these points in our results. It will be remembered that the average score in the Control group with reference to which each subject got his relative standing, was about 9; in the Profs and WPA groups it was 12. It appears, then, that for the subjects whose scores were above the average in all three groups, the average score functioned differently from the way it did in the subjects whose performance was below the average.

It is our hypothesis that the average score for those who are below it, functions as a particular concrete standard to be reached, while for those who are above it, the standard still continues to exist, but does not exercise a specific pull on the subject's estimates. Estimates can be made from the vantage point of greater self-confidence and on a somewhat freer basis. Our interview material reveals pretty clearly the differences in the meaning of the standard for the subjects whose performance was above and below it. It is fairly characteristic of our subjects that those below the average tend to mention it as a factor governing the formations of their own estimates, while those above the average rely more on their own previous performance and self-evaluation.

For example, we quote the response of two of our women subjects in performance category IV to the question: "How did you arrive at your estimate?"

> J. K., female, WPA, av. $P$[6] = 14.00: "Based on the time before, I wanted one point more than I had gotten."

---

[5] It might be said, in explanation, that since there was a maximum score of 20, those subjects in the highest performance category were mechanically limited in the range of their estimates. It must be noted, however, that the highest average performance was 14, leaving a possible average range of 6 points for aspirations. The D-scores in our lowest performance category never exceeded 6 points.

[6] $P$ refers to performance-score.

S. Z. R., female, WPA, av. $P = 14.50$: "I put down 18 or 20 because, I guess I was under the influence of my other intelligence tests. I have an IQ of 180, you know (sic!). I thought I could do it. . . . Somehow, I didn't think the papers were marked accurately. I felt I should have gotten 19 or 20."

The comments of another woman subject whose average performance was 13.14 in the Profs group is also revealing:

The first time, I did use the 12 as a standard. I got 11 and wanted to get at least 12. But afterwards, I discarded it, because I got way above 12 and I used my own Performance after that.

These are in contrast to the protocols of subjects below average, whose estimates were based on the average. Thus, a woman subject in the Profs group whose average $P$ is 4, says:

E. R., female, Profs, av. $P = 4.00$: "I tried to reach the average. I was below it. I put down 12 as average. I hoped and tried to hit it. I really didn't expect to get 12. I sort of felt I wouldn't, but I'd try. I felt I wanted to get average." (These comments are particularly interesting in view of the subject's ideas concerning the purpose of the experiment: "You were supposed to hit approximately your goal.")

Here is the answer of a subject in the WPA group (av. $P. = 7.43$).

J. O'M., male, WPA, av. $P = 7.43$: "I didn't know what kind of WPA workers they were—of normal intelligence or ditch-diggers. . . . So I thought I ought to do as well or better, and after I didn't do *13* I thought, well, I won't pull it down—maybe I'll hit it one of these days." (Note that 13 is 1 point above 'average.')

O. S., male, WPA, av. $P = 9.43$: "Based it on previous performance, more or less.[7] My average $P$ was about 10. Av. estimate—about 15." Q. "Why 15?" "I wanted to get a 15. I didn't feel very well getting a 10. The same psychology as everyone else. I felt I was better than WPA. . . ."

In some cases, although the subject spoke of basing his estimates on his previous performance, his answer was more revealing than he knew.

M. O., Male, WPA, av. $P = 7$, answers the simple question: "What was your general performance" by saying: "I don't think I ever reached 12." At the end of the interview, in answer to the question: "Were you influenced in any way by the figure 12?" "I judged myself according to the 12. If I had reached

[7] It is characteristic of subjects to mention previous $P$ in first answer to the question: "On what did you base your estimate?" But the below average go on to speak of the external standard (for example, this subject); others above the average, continue to discuss their own previous performance history.

12, I might have considered myself successful." Q. "What was your general estimate?" "About 16." Q. "How come?" "I refused to accept the fact that I couldn't answer the questions as I thought I should. You didn't allow very much time." Q. "Why 16?" "16 is better than average."

One male subject, whose performance was around 12, but who was sure that WPA workers 'could never make a score of 12' puts the point most sharply when he says:

R. T. E., male, WPA, av. $P = 12$; "Most of my thoughts were directed toward myself not doing as well as they did, rather than that they were doing better than they (the WPA workers) should have."

Changes in the role of the external standard are also indicated by the typical comments of a subject who, in the first few trials, made a lower score than the average, but later 'hit' the average and did even better.

T. S., male, WPA, av. $P = 12.33$; "At first I was told the WPA score was 12. Naturally, I wanted to rise above that. And so, in the beginning, my estimates were around 16. Then, once I got 16, and from then on my estimates were 17, because I wanted to improve on my own performance. My purpose in the experiment was nothing other than improving on the WPA or improving on my own previous score."

The largest mean D-scores in all 3 groups are found in the lowest performance categories. This can be seen from Tables IV and V, and it is also apparent from these tables that the differences in D-score between the lowest and highest performance groups either approach being or are statistically significant. Our results here offer independent confirmation of Anderson and Brandt's finding (1) that children in the lowest performance quartile had the highest goals, although the reverse relationship, in which the highest performance groups give lower aspirations than their preceding performance score, we do not confirm. The difference between our results may depend on the small number of cases in Anderson and Brandt's study.

Our protocols, again, show differences in the meaning of the estimate for low and high performance groups. The following two cases illustrate the fact that those in the lowest performance categories are pulled upwards in their estimates by virtue of the presented standard and of their low performance. The meaning of the estimate for these subjects seems

TABLE IV

MEANS, DIFFERENCES AND RELIABILITIES OF PERFORMANCE, ASPIRATION AND D-SCORES, ARRANGED ACCORDING TO PERFORMANCE CATEGORIES

| | Performance | | | | Aspiration | | | | *D*-Score | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | $\sigma$ | *D* | $D/\sigma D$ | Mean | $\sigma$ | *D* | $D/\sigma D$ | Mean | $\sigma$ | *D* | $D/\sigma D$ |
| *Control* | | | | | | | | | | | | |
| I. 5–8 (27) | 6.65 | 1.04 | | | 9.07 | 3.11 | | | 2.65 | 3.30 | | |
| | | | 2.06 | 9.36 | | | 0.97 | 1.23 | | | 1.87 | 2.60 |
| II. 8–10(18) | 8.71 | 0.28 | | | 10.04 | 2.17 | | | 0.78 | 1.50 | | |
| | | | 2.49 | 14.64 | | | 2.87 | 4.35 | | | 0.92 | 1.88 |
| III. 10–12(14) | 11.20 | 0.50 | | | 12.91 | 1.52 | | | 1.70 | 1.27 | | |
| | | | 2.61 | 5.67 | | | 1.10 | 1.57 | | | 1.46 | 1.52 |
| IV. 12– (7) | 13.81 | 1.17 | | | 14.01 | 1.52 | | | 0.24 | 2.37 | | |
| | | I–IV | 7.16 | 14.92 | | I–IV | 4.94 | 6.02 | | I–IV | 2.41 | 2.21 |
| *Prof* | | | | | | | | | | | | |
| I. 5–8 (18) | 6.01 | 1.23 | | | 10.59 | 1.10 | | | 4.15 | 0.99 | | |
| | | | 2.78 | 7.51 | | | 0.39 | 0.81 | | | 1.45 | 3.72 |
| II. 8–10(15) | 8.79 | 0.49 | | | 10.98 | 1.57 | | | 2.70 | 1.24 | | |
| | | | 2.14 | 10.70 | | | 0.71 | 1.78 | | | 1.39 | 3.86 |
| III. 10–12(17) | 10.93 | 0.62 | | | 11.69 | 0.18 | | | 1.31 | 0.68 | | |
| | | | 3.23 | 8.28 | | | 3.81 | 14.65 | | | .07 | .02 |
| IV. 12– (13) | 14.16 | 0.77 | | | 15.50 | 0.99 | | | 1.38 | 1.11 | | |
| | | I–IV | 8.15 | 16.30 | | I–IV | 4.91 | 13.27 | | I–IV | 2.77 | 7.10 |
| *WPA* | | | | | | | | | | | | |
| I. 5–8 (11) | 7.08 | 0.68 | | | 11.67 | 2.42 | | | 4.48 | 2.15 | | |
| | | | 2.07 | 7.39 | | | 0.80 | 0.77 | | | 2.21 | 1.97 |
| II. 8–10(13) | 9.15 | 0.75 | | | 12.47 | 2.66 | | | 2.27 | 3.35 | | |
| | | | 1.69 | 7.04 | | | 1.15 | 1.38 | | | 0.44 | 0.44 |
| III. 10–12(10) | 10.84 | 0.45 | | | 13.62 | 1.47 | | | 2.71 | 1.35 | | |
| | | | 1.77 | 4.02 | | | 2.30 | 3.38 | | | 0.23 | 0.45 |
| IV. 12– (11) | 12.61 | 0.89 | | | 15.92 | 1.89 | | | 2.48 | 1.25 | | |
| | | I–IV | 5.53 | 12.02 | | I–IV | 4.25 | 4.62 | | I–IV | 2.00 | 2.67 |

TABLE V

SHOWING MEAN D-SCORES FOR 4 PERFORMANCE GROUPS UNDER ALL CONDITIONS COMBINED, TOGETHER WITH $\sigma$'S, DIFFERENCES AND $D/\sigma D$'S

| Group | N | Mean *D*-Score | $\sigma$ | Difference | $D/\sigma D$ |
|---|---|---|---|---|---|
| I | (56) 5– 8 | 3.49 | 2.67 | | |
| | | | | 1.66 | 3.39 |
| II | (46) 8–10 | 1.83 | 2.30 | | |
| | | | | 0.07 | 0.18 |
| III | (47) 10–12 | 1.90 | 1.28 | | |
| | | | | 0.39 | 1.05 |
| IV | (31) 12 | 1.51 | 1.77 | | |
| | | | Diff. I–IV | 1.98 | 4.13 |

to be colored by hope rather than by a more sober estimate of possibilities. Thus, a woman subject in the Profs group, whose performance is low, both objectively and by her own subjective standards (she is ordinarily an 'A' student), says:

S. W., female, Profs, av. $P$ = 9.67: "I based my estimate on what I thought I really could do."

Or another woman subject in the Profs group whose score is 5.67, says:

> A. A., female, Profs, av. $P$ = 5.67: "I wasn't at all positive I'd get it—but I tried to. I was sort of boosting myself—boosting my ego."

E. R.'s protocol, quoted above, expresses a similar point.

Our interviews also reveal a difference in the meaning of performance and external standard between members of the WPA and Profs group. For the WPA group, it is necessary to do better than 12; for the Profs group it is enough to reach 12. It follows from this difference that it is easier for a subject to justify a score below 12 in the Profs group ("Oh, well, Professors only got 12"), while in the WPA group, a score below 12 cannot so easily be justified.

These qualitative results give meaning to Table VI which shows differences among our three conditions in mean D-

TABLE VI

SHOWING DIFFERENCE BETWEEN MEAN D-SCORES UNDER 3 CONDITIONS BY $P$ GROUPS

| | Difference | $D/\sigma D$ |
|---|---|---|
| Control-Prof | | |
| 5– 8 | 1.50 | 3.75 |
| 8–10 | 1.92 | 11.29 |
| 10–12 | 0.39 | 1.95 |
| 12 | 1.14 | 2.00 |
| Control-WPA | | |
| 5– 8 | 1.83 | 4.57 |
| 8–10 | 1.49 | 6.77 |
| 10–12 | 1.01 | 5.05 |
| 12 | 2.24 | 3.72 |
| WPA-Prof | | |
| 5– 8 | 0.33 | 0.82 |
| 8–10 | 0.43 | 1.80 |
| 10–12 | 1.40 | 7.00 |
| 12 | 1.10 | 3.66 |

score, by performance categories. The fact that performance category I (5–8) in the control condition felt itself closer to the obtained average of that group than did the I categories in the other two groups is shown by the fact that the mean D-score of category I is lowest in the control group (Table IV). The difference in D-score between control and experimental

groups in the I category are all reliable. There is no such difference between the I categories of the WPA and Profs. In a similar way, one can understand that reliable differences exist between the II (8–10) categories comparing Control-Profs and Control-WPA, and the lack of such differences between WPA and Profs.

Since the IV category (12 and above) was psychologically more distant from the 'average' in the Control group than in the other two, one would also expect reliable differences in mean D-scores between Control-Profs and Control-WPA. The D-scores of the subjects who scored 12 and above in the WPA group are higher than those of the comparable performance category in the Control conditions and the Profs condition. The actual performance scores, one can see, take on psychological meaning in accordance with their distance from the approved or desirable standard. To put it another way, the high-scoring subjects in the WPA and Profs groups were evaluating their performance against a higher standard. That one finds reliable differences in D-scores only between Control and WPA groups in the IV category can perhaps be understood from the fact that the performance above 12 was the least satisfactory for the group competing with 'WPA workers.' It was *most* necessary for subjects in this group to do *better* than 12.

This study has thus revealed some of the ways in which an external standard may operate to influence the nature of the subject's expressed goals. If we use as a criterion of change the difference in size of the mean discrepancy score between control and experimental groups, we see that knowledge of another group's performance does change the subject's own goal. The average discrepancy score of the control group was lower than that of the experimental groups. The change is a function, also, of the subject's actual distance in performance from the 'average.' If is for this reason that we obtained differences in mean discrepancy scores as a function of the size of the performance score. Our interviews indicate further that the relatively high discrepancy score which the lowest performance groups were forced to maintain led these

subjects to imbue their estimates with a factor of hope in addition to expectation. Finally, the social meaning attached to the 'average' increases or decreases the subject's psychological distance from it. Concretely, a score of 7, for example, was nearer the 'Professors' average of 12 than an objectively identical performance made by a member of the 'WPA' group.

*Summary*

1. The differences in performance among our 4 categories are reliable, under all three conditions.

2. Differences between the aspiration scores of the extreme performance categories are reliable under all three conditions.

3. The 'crucial' points in performance in each of the three groups, *i.e.*, the performance scores hovering around the standard, are the only points within the 4 performance categories which yield reliable differences in aspiration.

4. The D-scores of the performance groups operating under different standards are reliably different.

(Manuscript received April 25, 1940)

References

1. Anderson, H. H., and Brandt, H. F., Study of motivation involving self-announced goals of fifth grade children and the concept of level of aspiration, *J. soc. Psychol.*, 1939, **10**, 209–232.
2. Chapman, D. W., and Volkmann, J., A social determinant of level of aspiration, *J. abnorm. (soc.) Psychol.*, 1939, **34**, 225–238.
3. Gardner, J., Level of aspiration in response to a prearranged sequence of scores, *J. exp. Psychol.*, 1939, **25**, 601–621.
4. Gould, R., The effects of social pressure on aspiration level. To be published.
5. Hertzman, M., and Festinger, L., Group influences on the level of aspiration. In press.

# SHIFTS IN EXPLICIT GOALS IN A LEVEL OF ASPIRATION EXPERIMENT

BY MAX HERTZMAN

*College of the City of New York*

AND

LEON FESTINGER

*University of Iowa*

## *A.* Problem and Literature

The introduction of the difference score in level of aspiration experiments by Frank (**2**), while it filled a need by quantifying the results of aspiration experiments, brought a number of difficult problems with it.[1] According to Frank, the difference score is determined by factors which are permanent in the individual's total economy and consequently it should be manifested in a similar fashion regardless of the tasks employed to measure it. On the other hand, one may look at the difference score as dependent on a variety of factors not all of which are operating at the same time or in all tasks. This may make for a variability in the difference scores of the same subjects obtained on a number of tasks performed under different conditions.

Using 3 different tasks, and employing small groups of subjects, Frank (**2**) obtained medium sized correlations among the difference scores derived from these tasks. Frank's results and the subsequent findings of Gould (**4**) and Gardner (**3**), the former obtaining low correlations and the latter corroborating Frank's work have focused some attention upon the problem of the generality of the difference score.

Through a detailed analysis of interview material, Gould demonstrated that the same objective difference score could

[1] The difference score has generally been obtained by subtracting the individual's actual performance from his estimate of the level of his future performance. If his estimates are higher than his performance, the difference score will be positive and if lower a negative difference score will be obtained.

serve various purposes for different individuals. While such a state of affairs would not necessarily affect the consistency of the subjects' difference scores in a variety of tasks, it was shown that the subjects' aspiration estimates were to a great extent situational. The same underlying factors could be operating in each of the test situations for some subjects, but specific features of the situation could shape the aspiration estimate and consequently diversify the difference scores.

If such variability is usual, there are two important inferences which may be drawn. One would be that changes in difference scores could be effected by the appropriate manipulation of the experimental situation. In one of their experiments, significant changes in aspiration level were obtained by Chapman and Volkmann (1) by altering features of the experimental conditions. Unfortunately these changes were not studied in terms of the underlying attitudes of the subjects involved so their significance for the individuals of the study was not fully explained. A second, and perhaps more significant consequence, might be that changes in aspiration level and in difference scores brought about by specially designed experimental conditions could clarify the mechanisms at work in determining the individual's strivings. At any rate, experiments that would be successful in changing the aspiration level in relation to the performance level could contribute information concerning the operation of field factors in determining difference scores.

The present experiment was designed to introduce changes in an experimental situation after aspiration levels and difference scores have been determined on subjects who had no information other than their own performance and difference scores. The changes introduced were given to the subject in the form of the performance and estimates of comparable groups of subjects. In order to create as similar a psychological situation as possible for each of the subjects, the reported data were fictitious and arranged so as to put each subject in the same relationship with respect to the group. Thus the average performance of the group was always made equal to the average performance of the individual to whom

it was reported while the estimate of the group was fixed in the direction opposite to that of the individual. It was felt that of the various relationships that were possible between an individual and his group with respect to performance and aspiration this was one that was calculated to have a strong influence on the behavior of the subjects. It was hoped that knowledge concerning the character of the factors contributing to the difference score as well as information about the more permanent mechanisms and tendencies effective in an aspiration situation could be uncovered.

## *B.* Procedure

1. *Subjects.*—The subjects were all male undergraduates of the College of the City of New York. Most of them were in their sophomore or junior year. The total number of subjects used was 20.

2. *Tasks.*—Two tasks which satisfied our requirements were chosen. Tasks where a minimum of learning occurred were needed, that is, where performance remained essentially on the same level throughout. The two tasks selected were:

(*a*) *Synonym test*—15 lists of 24 words each were used. One of the lists was utilized for practise in the first session and each of the other lists was treated as a separate trial. The subject was asked to write after as many of the words from the list as he could words meaning the same as the printed word on the sheet. For 5 of the trials the subject was instructed to write synonyms beginning with the letter A. For another 5 trials the synonyms had to begin with the letter B, and for the other 5 trials, C was the letter to start with. The A, B, and C lists were presented alternately. The subject was allowed one and one-half minutes to complete the list, the trial being scored immediately. The subject was then told his score and was asked, "What do you intend to do next time?"

(*b*) *Information test*—17 lists of 22 questions each were used. One of the lists was used for practise as in the case of the synonym series, and each of the others was treated as a separate trial. The subject was told to answer as many of the questions as he could in two minutes. The test contained questions of the following type:

Where is Cathay?
The Goddess born from the head of Jupiter was . . .?
What year was Huey Long assassinated?
Who killed Macbeth?
What is a "Horse Opera"?

The test was scored immediately, the subject was acquainted with the score and was then asked, "What do you intend to do next time?" Average difference scores were obtained from these data by subtracting the average performance of each subject from his average estimate.

3. *Conduct of the Experiment.*—The experiment was divided into two sessions. In the first session the practise list and the first 7 synonym lists were given to the subject, to be followed by practise in the information test and the first 8 information lists. The second session took place between 8 and 10 days after the first one. In

the second situation, the subjects were retested with the last 7 synonym tests and last 8 information lists. The experimental factor was introduced in the second session in the following manner. In addition to telling the subject his score after each trial, he was also told what the average performance and the average estimate was for a fictitious group of 50 subjects before making his own estimate for the subsequent trial. The information he received concerning the group scores was adapted to each individual in the following manner. The figures given to each subject as those of the group were based on his own performance and his average difference score in the first session. The performance of the group as reported to the subject was identical, on the average, with the subject's performance on that task in Situation I. Thus, for example, if the subject scored an average of 12 in the information tests in Situation I, the mean of the scores reported for the group on the information tests in Situation II would also be 12. The scores for each separate trial were varied by the experimenter around a mean of 12 so as to make the situation realistic since no actual subject received the same performance scores on each trial of a given task. The estimates reported for the group were dependent on the average difference score for the individual in Situation I. The group estimates were the estimates of the individual in Situation I plus or minus 4 points depending on whether the average difference score of the individual was negative or positive. Thus, for example, if a subject had an average difference score in Situation I of plus 1.8, the estimate reported to this individual as the average estimate for the group would be 4 points below his own estimate in Situation I, and consequently equal to minus 2.2. And of course, if in Situation I the subject had an average difference score of minus 1.8 then the group estimate presented to him would be plus 2.2. Four points was chosen as the extent of the shift for several reasons. The shift had to be a medium sized one. If the shift were too small it might be overlooked by the subject, or even if it were not overlooked it might not have any effect upon him. On the other hand, the shift could not be too large or the subject might be skeptical as to the validity of the scores presented to him as those of the group. Even if he did not suspect them the large difference might remove all sense of reality from the situation and it again would not have any effect on the subject. This latter point is brought out very strongly by Hoppe (6). He found that fictitious results had meaning for the subjects and induced feelings of success and failure only when they hovered about the subject's aspiration level. When, however, the falsified results given to the subject differed too much, either above or below, from the aspiration level, these results were regarded as fictitious, or else they lost all meaning for the subject.

To summarize, Situation II was controlled so that the performance of the group would be about equal to the performance of the individual while the estimates of the group would be about 4 points in the opposite direction to the trend evidenced by the subject in Situation I.[2]

4. *Scores Obtained.*—For each subject the average performance, average aspiration, and average difference score was obtained for the synonym and information tasks in both Situation I and II. In addition, the difference between the average difference

[2] In all cases except one, the group estimates were varied in the same direction in both the information tests and in the synonym tests. That is, if the shift in the group estimate was downwards for the group information tests, it would also be shifted downwards for the synonym series. In one case, however, this is not so. This subject had a negative score for the synonym lists and a positive one for the information and so the group estimates were shifted accordingly.

score in Situation I and the average difference score for the same subject in Situation II was calculated for each task. This measure was called the change score and was given a plus sign if the change was in the direction of the estimate of the fictitious group, and was given a minus sign if it was contrary to the direction of that group's estimate. Thus, if the reported estimates for the group were 4 units below the subject's estimation in Situation I, it was expected that his shift would be in a downward direction and such a shift was regarded as positive. If the reported estimates for the group were shifted upwards from the subject's estimates in Situation I, then it was expected that the subject would shift his estimates upwards and such a change was also regarded as a positive one.

5. *Interview.*—At the end of the second session an interview lasting about an hour was given to each subject. This interview attempted to cover most of the relevant points influencing the subject's estimates and the changes in his estimates when they occurred. The following points were covered:

(*a*) The original aim of the estimate and how it was changed in Situation II.

(*b*) The reaction of the subject to the performance of the group.

(*c*) The reaction of the subject to the estimates of the group.

(*d*) The reaction of the subject to the difference between the group scores and the group estimates.

(*e*) The reaction of the subject to feelings of failure and success and causes of feelings of failure and success in both situations.

## C. Results

### 1. *Statistical Evaluation of the Results*

The means and standard deviations of performance scores, difference scores, and change scores are to be found in Table 1. In this table, item 9 is the same as item 1 since the scores re-

TABLE 1

Means and Standard Deviations of Performance, Difference Scores, and Change Scores

| | Mean | S.D. |
|---|---|---|
| 1. Perf. Inf. in Sit. I | 10.36 | 3.80 |
| 2. Perf. Syn. in Sit. I | 8.48 | 2.78 |
| 3. Perf. Inf. in Sit. II | 11.17 | 3.94 |
| 4. Perf. Syn. in Sit. II | 9.74 | 3.65 |
| 5. Diff. Score Inf. in Sit. I | 1.37 | 1.30 |
| 6. Diff. Score Syn. in Sit. I | 1.06 | 1.10 |
| 7. Diff. Score Inf. in Sit. II | .32 | 1.32 |
| 8. Diff. Score Syn. in Sit. II | .34 | 1.26 |
| 9. Group Perf. in Inf. | 10.46 | 3.80 |
| 10. Group Perf. in Syn. | 8.48 | 2.78 |
| 11. Group Diff. Score in Inf. | −.46 | 3.01 |
| 12. Group Diff. Score in Syn. | −.68 | 3.07 |
| 13. Change Scores in Inf. | 1.22 | 1.19 |
| 14. Change Scores in Syn. | 1.19 | 1.46 |

ported for the group were the individual's performance in Situation I. The slight difference between these two scores is due to an adjustment that was nescessary in one case with a particularly low performance score. Item 10 is the same as item 2. We can see from items 5 and 6 that the original tendency of the subjects was to keep their estimates above their performance. We can also see from items 7 and 8 that a marked change occurred in Situation II and that the estimates, while still tending to be above the performance, approximated the performance more closely than they had done in Situation I. Since the tendency on the part of most individuals was to overestimate, the difference score for the group, is of course negative, as the estimates of the group were varied in the direction opposite to the estimates of the subjects. The average of the change scores for both the information series and the synonym series was positive, indicating that the general trend was to change the aspiration estimate in the direction of the group's estimate.[3]

TABLE 2

EVALUATION OF THE DIFFERENCES BETWEEN SOME MEANS IN TABLE 1

| | $Mean_1$ | $Mean_2$ | Diff. | S.E.$_{Diff.}$ | Ratio (D/S.E.$_D$) |
|---|---|---|---|---|---|
| Perf. I; Perf. II; Syn. | 8.48 | 9.74 | 1.26 | 1.06 | 1.19 |
| Perf. I; Perf. II; Inf. | 10.36 | 11.17 | .81 | 1.25 | .65 |
| G.D.S.;[1] I.D.S.; Syn. | .32 | −.46 | .78 | .75 | 1.04 |
| G.D.S.; I.D.S.; Inf. | .34 | −.68 | 1.02 | .76 | 1.34 |
| G.D.S.; I.D.S.; Both | .33 | −.57 | .90 | .52 | 1.73 |

[1] G.D.S. means group difference score; I.D.S. means individual difference score.

In Table 2, the differences between the means of selected performance scores and difference scores are evaluated. In Table 3, the significance of the change scores is determined.[4]

[3] When questioned as to whether they were consciously influenced by the group data, seven of the subjects said that they were while the remaining thirteen disclaimed a conscious awareness of any influence. When the results of the two groups of subjects were treated separately the same pattern such as is shown in the data of the total group was found in each of the two sub-groups thus formed. While the influence of the experimental factor was more pronounced in the "conscious" group the difference between the averages of the two groups was not statistically reliable.

[4] Because the number of cases was less than thirty, $\sqrt{N-1}$ instead of $\sqrt{N}$ was used in calculating the standard errors (5).

TABLE 3

EVALUATION OF THE SIGNIFICANCE OF THE CHANGE SCORES

| | Mean Change | S.D. | S.E. Mean | Ratio $\left(\frac{\text{Mean Change}}{\text{S.E. Mean}}\right)$ |
|---|---|---|---|---|
| Change Score for Syn. . . . . | 1.19 | 1.46 | .31 | 3.84 |
| Change Score for Inf. . . . . | 1.22 | 1.19 | .27 | 4.52 |
| Change Score for Both. . . . . | 1.21 | 1.33 | .21 | 5.76 |

There were slight improvements in performance in the second situation as compared to the first but these differences were not reliable. For the synonyms, the ratio of the difference over the standard error of the difference is 1.19, while for the information series the ratio was only .65. Neither of these differences therefore can be regarded as statistically significant. The next three items in Table 2 represent the difference between the experimentally introduced group difference score and the individual difference scores. Similarly, none of the three ratios obtained here are found to be statistically significant.

We find significant ratios ranging from 3.84 to 5.76 in determining the significance of the change scores. The difference in performance between the two situations (which is not a reliable one) is a minor factor in accounting for the change in difference scores. The marked significance of these changes must be due largely to the experimental factor, which for these subjects consisted of being confronted with group estimates at variance with their own, and performances equal to their own. A further statistical corroboration of the conclusion that the individuals tended to make their estimates conform to the group estimates can be seen in the fact that the differences between the difference scores reported for the group and the individual difference scores in Situation II are not significant. In other words, the difference scores of the individuals tended to approach the difference scores reported for the group. The subjects on the whole, have changed their manner of estimating so as to tend to conform to the group, even though the group was opposed to their original tendencies.

While correlations based on as small a sample of subjects as 20, are not likely to be stable, the intercorrelations of the performance scores and the average difference scores presented in Table 4 are suggestive.[5] Most interesting is the fact that

TABLE 4

INTERCORRELATIONS OF PERFORMANCE SCORES[1] AND OF DIFFERENCE SCORES

| | Inf. I | Inf. II | Syn. I | Syn. II |
|---|---|---|---|---|
| Inf. I | | .93 | .56 | .36 |
| Inf. II | .44 | | .45 | .17 |
| Syn. I | .16 | −.09 | | .65 |
| Syn. II | −.02 | .38 | −.02 | |

[1] Correlations above the diagonal are between performance scores; those below the diagonal are between difference scores.

while the performance correlation between the two synonym tests is fairly high being equal to .65, the correlation between the average difference scores of these two tests is −.02. While the change of conditions has not affected the relative position in performance markedly it seems to have affected the way in which the subjects make their aspiration judgments. That the experimental conditions have brought about changes seems also to be reflected by the fact that correlations between the difference scores obtained from the synonym and information tests are higher when the tests occur in the same session than when they occur in different sessions. The latter result, however, may have been at least partially influenced by factors specific to the sessions in which the tests occur. The effect of the experimental factors is not uniform. A correlation of only .34 was obtained between the difference score reported for the group and the difference score for the subjects in Situation II for the information tests. A still lower correlation of .22 for these two variables in the synonym tests shows that there is no one to one correspondence between the size of the group estimate and the effect it had on the individual.

[5] The lowest correlation that is four times its own probable error is equal to .48 and the lowest correlation three times its own probable error is .39. (For the latter value the chances would be 98 in 100 that the correlation was significantly greater than zero.)

2. *Attitudes of the Subjects as Shown by the Interview Material*

(*a*) *Conscious Adaptation to the Group Estimate.*—Seven of the subjects admitted to adapting themselves consciously to the direction of the group estimates. Two of these subjects said that the estimates of the group became the norms for their estimates. Two other subjects said that they toned down their estimates because the group estimates were low. Another subject stated that the high estimates of the group created an atmosphere in which he could overestimate with equanimity. We see that all of these methods of adapting to the group estimates were in the direction of estimating in the same manner as the group did. Whether it was from the point of view of conforming to the group or competing with the group, the result was to shift the estimate of the subject in the same direction as the group estimates. There were two subjects who stated that the low group estimates influenced them to estimate higher. Neither of these two subjects was very clear as to the reason for this, however. These two are the only subjects who reacted in a way opposite to the group estimates, while admitting the influence of the group estimates.

(*b*) *Apparent Rejection of the Group Estimates.*—Many more of the subjects insisted that the estimates of the group did not have any influence on them of which they were aware. The methods of denying any influence of the group estimates upon them were so uniform that 11 of the 13 subjects who apparently rejected the group estimates insisted that they either thought the group estimates unrealistic or they completely ignored them. Of these 11 subjects, 5 later in the interview admitted that perhaps the group estimates created an atmosphere which influenced their own estimates but which did not influence them in a manner so direct as to make them consciously aware of the influence. The other 2 subjects under this heading denied that the group estimates influenced their own estimates in any way at all.

(*c*) *Nature of the Adaptation to Group Performance.*—In 15 of the 20 subjects, the estimates in the second situation were a direct function of the relation between the subjects'

scores and the scores of the group. The way the estimates were used in this relation was precisely the same for all 15 subjects. These subjects thought that they should be able to score above the group performance or else at about the average of the group performance. As a result of this, when the subjects scored below the group, their estimates were high, namely, above or at the group performance level. When the subjects scored above the group performance level, the striving was removed from their estimates and their estimates dropped below their performance. In the latter case, the estimates were usually kept between their own performance and the performance of the group, so as to reinforce feelings of success and to beat the group in as many ways as possible. By keeping their estimates between their own performance and the performance of the group they took only a small risk of not surpassing their estimates.

For these subjects (constituting 75 percent of the total number) the addition of the group's performance score has resulted in the introduction of a new factor in the total field situation. This new factor has become a goal which can be set without loss of prestige and it consequently has a marked effect in changing their aspiration estimates and difference scores. For such subjects, the aspiration estimate might thus be the consequence of a momentary striving situation rather than determined by basic goals.

(*d*) *Change of Goals in Situation II as a Consequence of Group Performance.*—The comparison with the group performance resulted in a change in the nature of the explicit goals of the subjects, whether their estimates were affected or not. The goals of the subjects were clarified through this comparison and meant more to them in addition to their becoming more standardized. While originally the goals of the subjects had been entirely independent of any outside influence, since the subjects knew nothing about what scores should be attained on the tests, now the goals of the subjects were in relation to the group performance. In other words, these subjects wanted either to beat the group or equal the performance level of the group. In this way the goals of the subjects be-

came more concrete. While before, the subjects' aspirations may have taken flights into unreality and may have been extremely high, in Situation II they were tied down to a specific factor, namely, the scores of the group. Whereas in the first situation feelings of success and failure had been experienced only in relation to the estimates of the subjects or with reference to the other goals which the subjects had, in Situation II feelings of success and failure were experienced mainly in relation to the group performance.

### *D.* Discussion

We have found reliable shifts in the aspiration estimates of our subjects. The meaning of these shifts was studied through data obtained from interviews concerning the first and second situation. From the quantitative data we discovered that most of our subjects were setting standards that were higher than their own achievement. We may present the following hypotheses in connection with the original strivings of the subjects. (1) That the high explicit strivings were determined by desires to do well in the task. (2) That the continued discrepancy between their performance and estimates was creating a tension in the subjects which could be relieved by the introduction of appropriate factors. The experimentally introduced factor brought about a change in the situation. Each subject discovered the group as a whole was doing no better than he was. As we know, the estimates of most of the subjects were reduced with the majority disclaiming that the low group estimates were factors influencing the reduction of their own estimates. The reduction of the estimates can be seen, however, in light of the two hypotheses stated above. The subject could relieve the tensions established through the discrepancy between the scores and estimates by reducing his own estimates. At the same time, he was not relinquishing a satisfactory goal as he discovered that in terms of the group performance his own score was not a bad one. What may have occurred for most of the subjects then, was a shift in explicit goals (which does not necessarily carry the implication with it that implicit goals had been

changed). The desire to do well still existed but a more stable criterion of good performance had been established. This enabled the new criterion to become a goal rather than the original high estimate established by the subject.

One might also ask why under these circumstances was the average of the group accepted by the subjects if they were desirous of making a good score. Would they not really aspire to do better than the average? A number of factors must be considered in evaluating the apparent acceptance of the group performance as a goal. We must remember that if the discrepancy between the performance and the estimates of the subjects created a state of tension they would more readily accept a passable but adequate goal which could relieve their tension than continue to set an unreachable goal. Moreover, the subjects' aims did not seem to be that of merely reaching the group performance but of doing a little better. This is shown by the fact that while the difference scores were reduced they still remained positive, as well as by the increased estimates that most subjects gave when they fell below the group performance on individual trials of a given test. Subjects scoring continually below the group, and thus being unable to compete with the group on the plane of performance, often raised their estimates to the level of the group's performance and so received some satisfaction from placing their estimates in competition with the performance of the group. For most of the subjects under the new circumstances, it was, however, possible to equal or exceed a goal that was satisfactory to the ego without bringing in a tension due to setting an explicit standard that was too high.

The group's estimate may be a factor in determining the lowering of the individual estimate in the second situation, but the manner and extent of its influence are not clear. While in the statistical results we find a definite change in the direction of the group estimates, yet in the interview 13 out of the 20 subjects flatly denied that the group estimates had any direct influence. The great majority of difference scores were positive and the lower estimates of the group could conceivably have created an atmosphere in which social

sanction was given to the lowering of the estimate. Tensions created through the attempt to maintain too high a standard could also be relieved by conforming with the group.

The results of this experiment show that the aspiration level can be greatly affected by the organization of the field that the subject finds himself in. General motives of avoiding failure and relieving tensions may be thought of as being strong determining factors in the subject's behavior. In the first situation, the establishment of high goals may help to avoid failure but it creates tensions. The second situation enables the subject to avoid failure by utilizing an acceptable substitute goal which relieves the tension created by a discrepancy between his strivings and achievement.

The data of this study would fit into the hypothesis that the expressed level of aspiration is the consequence of a number of factors entering into the specific situation that the subject is faced with and that with the alteration of the field situation the explicit aspiration could change considerably. Such findings would indicate that if there is a legitimate problem in connection with the generality of the aspiration level it is not to be solved by dealing with correlations between difference scores obtained in a variety of different situations. Rather, it must deal with investigations concerning the more permanent bases of the individual's strivings.

## *E.* Summary and Conclusions

1. Performance scores, aspiration estimates, and difference scores between their performances and aspirations were obtained on 20 male college students using several trials of a synonyms test and an information test.

2. In the first session of the experiment, positive difference scores were obtained for the vast majority of subjects.

3. In a second session, each subject was informed of the average performance and average estimate of a fictitious group of 50 fellow students at the end of each trial in each of the two tasks. Actually the performance level reported was equal to the subject's own performance obtained in the first session while the estimate of the group was in the direction

opposite to that of the subject, differing from his own estimate by about 4 points.

4. In the second session the difference scores of the subjects while remaining slightly positive were reliably reduced in magnitude.

5. Interviews conducted after the second session to ascertain the subject's perception of the two situations revealed the following:

(*a*) A majority of the subjects shifted their explicit goals from their own estimates to those of the group performance. The hypothesis was developed that the high level of aspiration created a tension which the introduction of a goal acceptable to the subject's ego relieved. (Since the fictitious group's average was just equal to the subject's own performance in the first session it could be reached or approached by most subjects no matter how poor their performance actually was. A number of subjects were able to exceed their earlier averages.)

(*b*) Most subjects denied being consciously influenced by the low estimates of the group. It is possible however that these low estimates created an atmosphere which granted social sanction to the lowering of the estimates.

6. It was concluded that the explicit level of aspiration could be varied considerably by changes in the specific features of an 'aspiration' situation. The problem of the generality of aspiration estimates could be investigated more fruitfully through studies of the factors underlying expressed level of aspiration than in terms of the difference scores or the aspiration estimates themselves.

(Manuscript received May 6, 1940)

## References

1. Chapman, D. W., and Volkmann, J., A social determinant of level of aspiration, *J. abn. (soc.) Psychol.*, 1939, **34**, 225–238.
2. Frank, J. D., Individual differences in certain aspects of the level of aspiration, *Amer. J. Psychol.*, 1935, **47**, 119–128.
3. Gardner, J. W., Aspiration-level in response to a prearranged sequence of scores, *J. exp. Psychol.*, 1939, **25**, 601–621.
4. Gould, R., An experimental analysis of 'level of aspiration,' *Genetic Psychol. Mono.*, 1939, **21**, 1–116.
5. Guilford, J. P., *Psychometric methods*, New York, McGraw Hill Book Co., 1936.
6. Hoppe, F., Erfolg und Misserfolg, *Psychol. Forsch.*, 1930, **14**, 1–63.

# AN ADAPTATION OF THE SMEDLEY HAND DYNAMOMETER FOR USE IN MEASURING VOLUNTARY FATIGUE [1]

BY R. FREDERICK BECKER AND HARRY N. GLICK

*Northwestern University Medical School*

*and*

*Massachusetts State College*

## I. The Ergographic and Dynamometric Method

In 1914, Glick (**7**), as reported in an unpublished thesis, devised an experimental apparatus for measuring voluntary muscular fatigue which, so far as is known, has not been duplicated since that time. Recently we have reconstructed a similar instrument which we believe eliminates some of the shortcomings commonly encountered in ergography and dynamometry.

It will be well to examine first some of the methods employed in the past in studying work and fatigue of simple muscle groups in the integrated organism. Of classic interest is Mosso's adaptation of the principles inherent in von Helmholz's myograph. In the construction of the ergograph for measuring work done in finger flexion against a constant weight, Mosso (**10**) attempted to isolate the working muscles as completely as possible so that no aid could be given them as they tired. The work curves were recorded on a revolving kymograph. An apparatus of this type is still widely used today, and has been the means of securing valuable information concerning muscular fatigue.

However, after the early enthusiastic investigations of Lombard (**8**), Mosso (**10**), Bergstrom (**2**) and Ash (**1**), interest in ergography waned for a time. It is quite impossible in the integrated organism to confine work to isolated members of a definite muscle group and expect these members to operate

[1] The apparatus was constructed in the Psychological Laboratories of the Massachusetts State College.

under constant conditions. Many investigators avoided the ergographic technique for this very reason. Yet there has not been substituted for study any other integrated activity entirely devoid of the variability inherent in the weight-lifting or spring-pulling of ergography. Renewed interest, on the part of the psychologists at least, is evidenced by the contributions of Crawley (**5**) 1926, Weinland (**11**) 1927, Manzer (**9**) 1927, and Yochelson (**12**) 1930. Using modifications of the Mosso apparatus these investigators studied the contractions of the single finger, of the hand muscles involved in grasping, of the flexors of the leg, and the large muscles of the trunk as used in rowing. None of these investigators are as interested in isolating the working muscles involved as were the earlier workers. In fact, at the present time, it is still next to impossible to study a simple muscle or muscle group in the integrated organism with the expectation of constant conditions and non-interference of other closely related muscle groups in spite of stops, straps, clamps and other devices.

Furthermore, there are other objections to the use of the Mosso type of instrument. With the weight ergograph, if the maximal contractions are repeated every two seconds, at first the extent of movement is great, but after fifty to one hundred contractions the muscle can no longer raise the weight, and no mechanical work can be accomplished. The total work, therefore, according to Mosso, which a muscle can do before it is fatigued, is the product of the weight and the sum total of the extent of the separate lifts. There is some question as to the truth of this statement. When a muscle can no longer lift a weight of three kilograms and consequently under this condition does no mechanical work, it is, nevertheless, not entirely fatigued and can still function. Indeed, a weight of two kilograms may, at this point, be lifted nearly as high or even higher than the initial lifts at three kilograms. Moreover, when a muscle pulls against a weight it cannot lift, considerable physiological work is being done even though no mechanical work is evident.

In addition, weight ergograms are hardly comparable even when two individuals pull the same weight until their finger

muscles are exhausted. Individuals have their own characteristic maximal pulls. If a man is set to work who can just raise three kilograms to its highest level once, certainly the rest of the curve will show a decided downward trend. If another individual is set to do the same task but who can easily lift five kilograms, his curve will show a long series of lifts of even height before there is any indication of a drop. In fact, if one examines the usual weight ergogram, there is nothing much there in the way of a fatigue curve at all. Usually there is a long series of lifts of equal height and about two or three decreasing lifts at the end. The end drop in almost every case exhibits a decided, sudden dropping-off which one would hardly expect of muscular fatigue. The weight ergograph has a further disadvantage common to all isotonic measurements. The changes in extent of movement from the early contractions to those at the end cause a corresponding change in muscular nutrition so that the condition of the muscle is wholly changed during the work period.

It was such objections as the foregoing which led Franz (**6**), Cattell (**4**), Binet and Vaschide (**3**) and others to devise their spring ergographs. By pulling against a spring of the proper tension the least as well as the maximal contraction of each individual could be recorded. There was also less chance for inertia to alter the height of contraction as compared with weight ergography. The very shape of the curve itself was more indicative of fatigue than that of the usual weight ergogram.

In spite of these advantages one must also recognize the limitations of such an instrument. The tacit assumption made by users of the spring ergograph, that the tension of the spring is immaterial provided the tension constant is sufficient enough to register maximum effort each time a contraction is made, is not entirely free from question. As Franz (**6**) demonstrated so well, if, after a series of contractions against the force of a heavy spring, a new series was begun on a lighter spring, the work done with the lighter spring was always greater than that done with the heavy spring. In addition, whenever a change was made from a heavy to a light spring

there was a decrease in the amount of tension overcome. In another experiment, work with a medium spring was alternated with work upon a strong spring. Sixty-six alternate work periods were run with ten-second rest periods between changes of springs. All results were the same; both the work accomplished and the tension overcome were different when different springs were used. In regard to work done, it was found that some subjects did more work against the medium spring. The amount of tension overcome was always greatest against the heavy spring. It would seem that a change in tension in spring ergography creates a condition similar to a change in weights on the Mosso instrument. With the spring ergograph, likewise, the condition of the muscle is markedly altered during the work period. We are dealing, however, with an oxotonic rather than isotonic effect in this instance.

Of recent years, one finds little reference to the use of the hand dynamometer as an instrument for measuring voluntary muscular fatigue though it was used quite extensively in the past. An early modification of the hand dynamometer for the study of fatigue was reported by Cattell (4) in 1897. There is no detailed description of this 'ergometer.' It is simply disclosed that an instrument was fitted to write upon a kymograph and the pressure of the thumb and forefinger as well as the movement of any finger could be measured.

In these early reports two methods were commonly employed—the method of sustained contraction, and the method of separate contractions. In the former case a metronome would be set to beat once a second. The subject was instructed at the signal, 'now,' to grip as forcibly as possible and maintain his grip with the utmost effort until told to relax at the end of one minute. The subject was to watch the hand on the dial and keep it as high as possible. The experimenter noted the initial reading and one at every fourth beat of the metronome thereafter. In the method of separate contractions the subject made a series of 16 grips as forcibly as possible at 4-second intervals. The experimenter jotted down the readings as previously.

The work of Binet and Vaschide (**3**) in 1897 is illustrative of studies of this sort. Using the separate contraction method upon a group of boys ages 10–13 and a group of young men age 18, each gripping 5 times alternately with each hand, these authors noted four types of endurance curves. These types represented: (*a*) a sudden drop, then fairly constant; (*b*) approximately stationary or constant type which they thought quite common; (*c*) a continuous, but gradual drop; (*d*) and a more or less definite rise—this type infrequent.

Outside of Glick's (**7**) unpublished thesis of 1914 it is hard to find any recent studies of note employing the hand dynamometer as a measuring device in relation to the work decrement.

In most of the early studies the instrument used was either the oval type which fitted in the palm of the hand or the larger Smedley Revised Dynamometer with or without pneumatic transmission. As Franz (**6**) indicates, the dynamometer has an advantage over the ergograph in that it employs an isometric spring. When a muscle pulls against an isometric spring the muscle does not shorten very much and the energy is converted mainly into tension. This very factor is a distinct improvement over the isotonic and oxotonic contractions encountered in ergography. In pulling against a weight or spring the muscle is continually lengthening and shortening, and these changes in extent of movement introduce another variable factor which should be eliminated if possible. When the movements are long, as in the beginning series of contractions on the weight ergograph, a considerable amount of blood and lymph is displaced about the muscle tissue. Circulation is increased and waste products are eliminated readily. When the movements become shorter, as toward the end of the series of lifts, the circulation becomes sluggish, wates products accumulate more rapidly and the condition of the muscle is wholly changed. With a constant extent of movement as in isometric contractions this difficulty is reduced to a minimum.

The obvious objection to the oval dynamometer is that the instrument can not be gripped in precisely the same manner

during consecutive contractions. A slight shifting about in the hand brings fresh muscles into play, and partially fatigued muscles have a chance to recuperate. Slipping due to perspiration is very apt to occur during a series of grips so that results are not strictly comparable. Likewise there is no allowance possible for variations in length and contour of hands with such an instrument.

In order to avoid some of these difficulties, Smedley invented a new type of dynamometer which is now commonly used more as an instrument to measure strength of grip than endurance or fatigue. However, it has been used for this latter purpose. There is less chance for the hand to slip in grasping the specially shaped handle. Moreover, the handle can be adjusted to fit almost any type of hand. As mentioned above, either the method of constant or intermittent grip was employed by the investigators using such an instrument to measure the work decrement. The intermittent method was preferred but the experimenter ran into more difficulty than the subject when this technique was used. With the metronome set at 60, the experimenter signaled the subject to grip on every fourth beat. The experimenter then had to read the dial, record the figure, reset the hand and get the subject ready for the next grip. Obviously the method was open to error in reading the instrument as well as allowing too much time for recuperation of the muscles between grips.

Smedley then produced his Revised Dynamometer fitted with an attachment for pneumatic transmission. A small, tightly-fitting piston head moves in a small aluminum tube. The tube is connected with a tambour by means of a slender rubber tubing. When the instrument is gripped the piston works in the tube tending to create a vacuum in proportion to the force exerted upon the dynamometer. Naturally, the vacuum effect is transmitted through the tubing to affect the diaphragm of the tambour. This, in turn, affects a stylus which records upon the drum of a kymograph in the usual way. An instrument of this type is marketed by the C. H. Stoelting Company at the present time.

Such an apparatus seemed to be the solution to the problem. Now all the experimenter had to do was to sit back and let the subject pull, and the tempo could be stepped up to once or twice a second, thus, eliminating recuperative effects. However, Glick (7) found that pneumatic transmission produced an inaccurate recording of the voluntary fatigue curve in three respects: (1) A quick grip tended to create a greater vacuum and hence affected the stylus more than a slow grip of the same force. Some subjects were inclined to grip quickly and let go slowly; others did just the reverse. Hence comparable records could not be obtained from two such adverse individuals. (2) A strong grip affected the recording more than a weaker one. For example, if a grip of 50 kilograms effected a throw of the stylus two inches, a grip of 20 kilograms may not have effected a throw of more than half an inch. The discrepancy is probably explained as an effect of inertia—a long throw carrying the stylus beyond the point where force alone could drive it. (3) Possibly the greatest defect of pneumatic recording was its failure to register holding effects. Few subjects relax completely between successive grips. That is, as fatigue develops, a certain amount of contracture of the flexors becomes noticeable. The hand of the dynamometer never comes back to the base line after each grip. Kymographic recordings give no hint as to the amount of holding the subject does. If the subject relaxes completely between grips at first and holds toward the end of the curve, his kymographic record will show a sudden drop. If the subject develops an early contracture but relaxes later, as some do, the end of the kymographic record will show but little drop. As far as the record of work is concerned, the two curves might be quite comparable, yet one would never suspect it from the kymographic recordings.

Since the method of pneumatic recording could not be relied upon, Glick (7) constructed an instrument of his own which was simply a Smedley Dynamometer with its three salient faults corrected. It is essentially this instrument which we have constructed in a recent study of voluntary muscular fatigue (to be reported later). There have been

some minor changes in order to facilitate recording. While the Smedley instrument is now only occasionally used to measure the strength of grip, with a bit of revamping, such as we are about to describe, it has inherent possibilities of becoming a good apparatus for measuring voluntary muscular fatigue.

## II. Construction of the Apparatus

The dial plate and marker-hand were removed from the usual Smedley Revised Hand Dynamometer, and the dial was replaced by a piece of sheet zinc 11 inches by 13 inches (*A*,

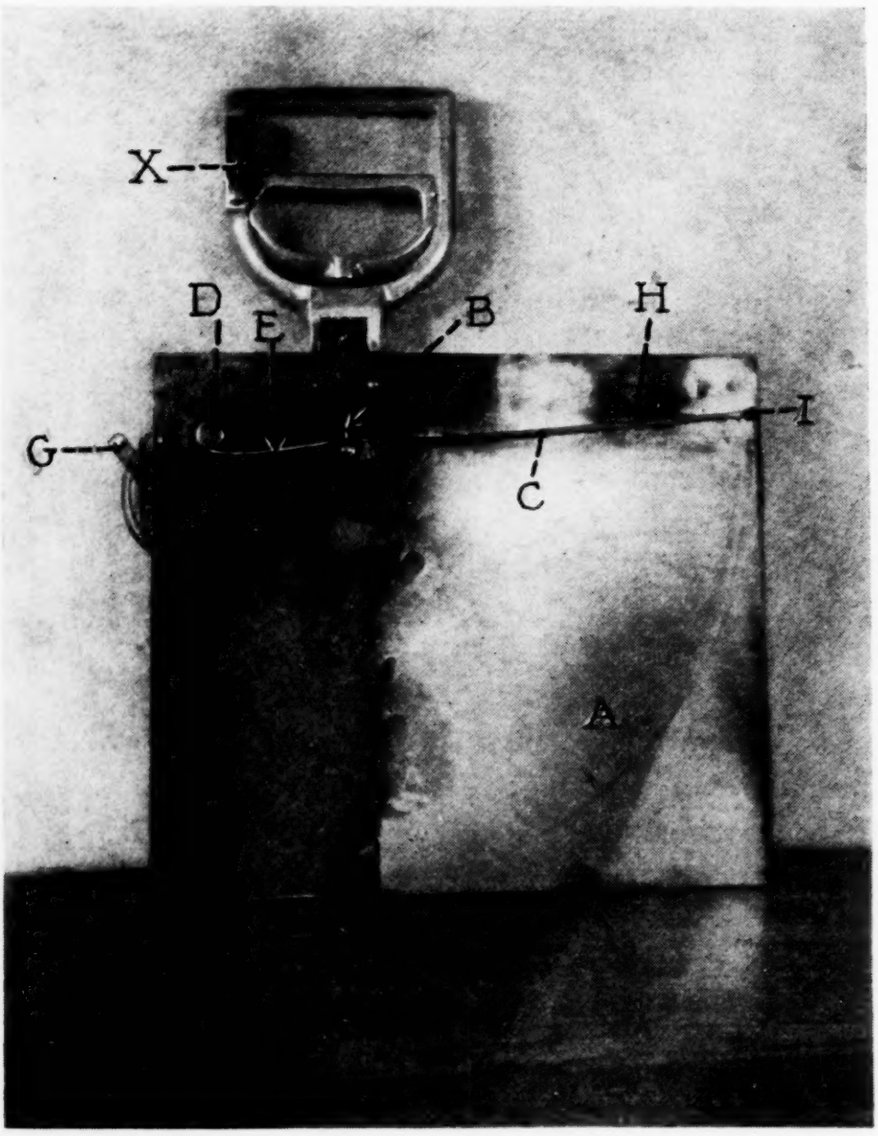


Fig. 1. Modified Smedley Hand Dynamometer. Front view.

Fig. 1) backed by cork boarding to hold it stiff (*A*, Fig. 2). An aluminum sheet would probably serve better. The instrument was not centered on the zinc sheet since about 8 inches to the right were needed for a prolonged hand while no

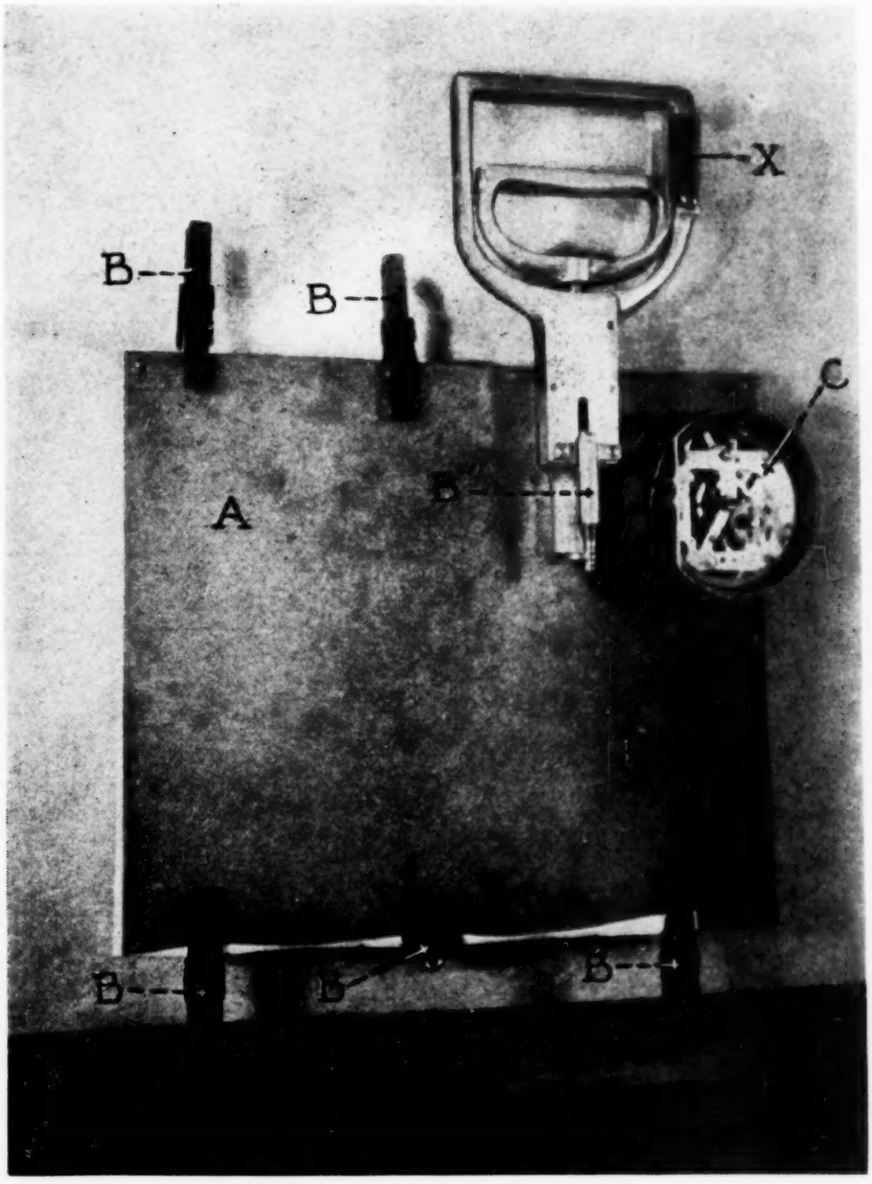


FIG. 2. Modified Smedley Hand Dynamometer. Rear view.

more than four inches were required to the left (see Fig. 1). The fixed hand was prolonged about 17 centimeters by a small, thin-walled, brass tube about 4 millimeters in diameter (*C*, Fig. 1). The tube may be either soldered or fastened to the fixed hand by small screws. The tube was sawed

open on the under side nearly to its full length—all except about 3 centimeters where it fastened to the hand. This slit in the tube was about 1 millimeter wide.

A small brass rod about 2 centimeters long was free to slide inside the prolonged tubular hand. A small, delicate spring stylus was soldered to the under side of the rod near its most distal end as it lay in the tube. The stylus projected through the slit in the under side of the tubular hand and made contact with the zinc surface. When the instrument was gripped, the long hand moved in an arc across the zinc surface for a distance proportional to the force exerted just as the original hand moved in an arc about the dial of the original Smedley instrument. If the surface of the zinc sheet were now covered with smoked kymograph paper cut to fit and fastened with spring clamps (*B*, Fig. 2), the stylus would leave a curved tracing in the soot indicating the extent of movement of the hand.

The problem was to have the stylus mark at a different place for each successive grip if the force of each was to be measured. The works of an alarm clock (*C*, Fig. 2) were fastened to the rear of the zinc sheet. The shaft for winding the main spring was extended through to the front and keyed to a spool (*D*, Fig. 1). The spool acted as a stem wind when turned in one direction. A small brass stud projected from one side of the spool about which a loop of thread (*E*, Fig. 1) could be slipped. The other end of the thread passed behind a peg guide (*B*, Fig. 1), through the tube, and fastened to the end of the stylus-rod. A lever (*G*, Fig. 1) started and stopped the alarm clock. The clock was then wound, the loop was attached to the stud on the spool, and the lever was raised. The thread then wound slowly around the spool and the stylus-rod was drawn up the tubular hand at a slow rate making a new mark in the soot each time the subject gripped. See Fig. 3 for a typical record.

A bumper (*H*, Fig. 1) covered with soft felt was fastened to the zinc sheet for the hand to strike, for it often came back violently when a subject relaxed completely between grips. At the distal end of the tube (*I*, Fig. 1) a small rider was fast-

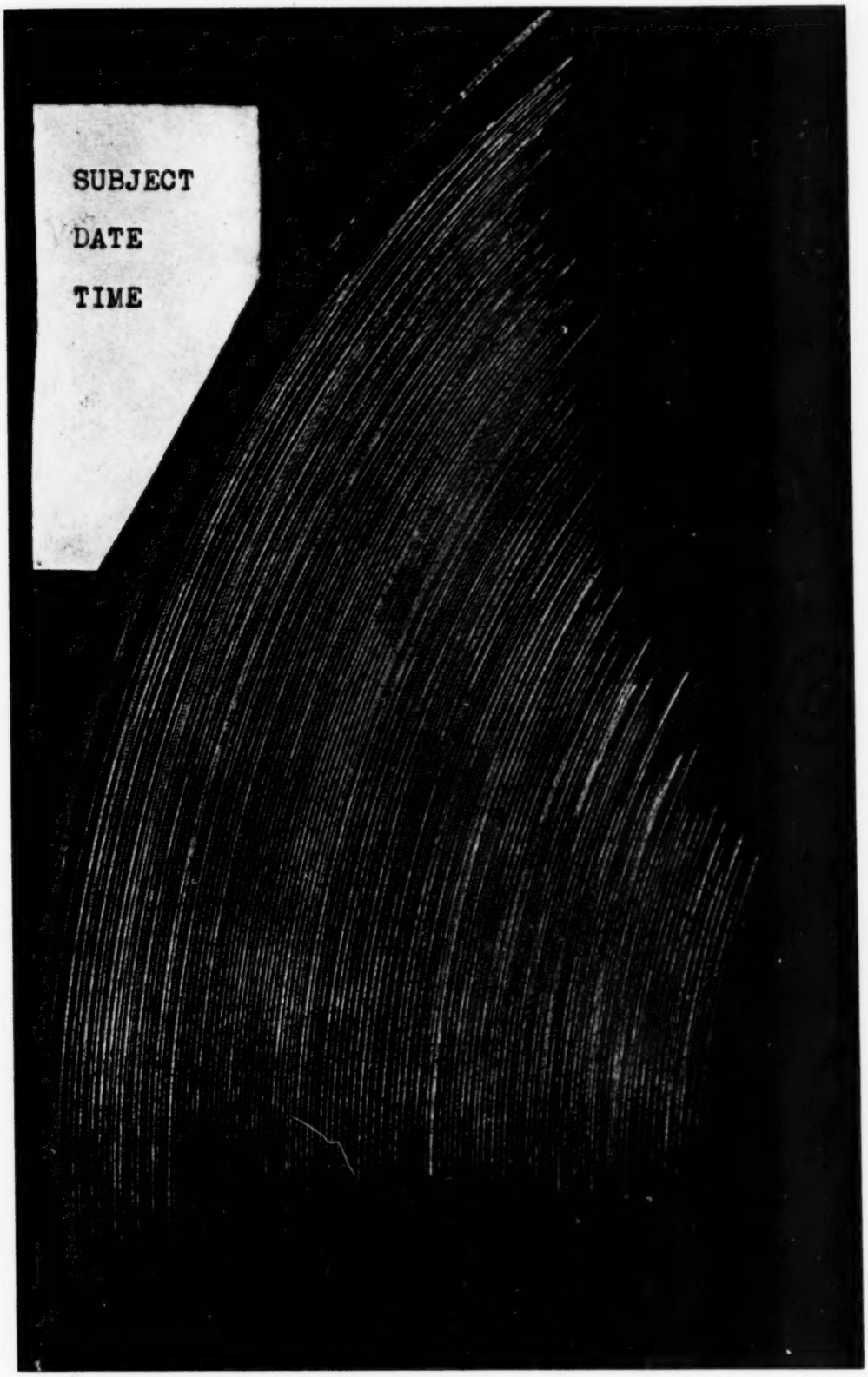


FIG. 3. Sample recording. Note the holding effects along the base line.

ened to run smoothly over the zinc surface, and to keep the long arm from dipping into the soot and spoiling a record. The path of the rider can be seen in Fig. 1 where it has marked the zinc surface by constant rubbing. In Fig. 2, '*B*'' indicates the aluminum attachment of the Smedley Revised Dynamometer to which the rubber tubing and tambour would be attached for kymographic recording. It is not necessary to use this attachment at all in our present instrument. The catch (*X*, Figs. 1 and 2) marks the point where the handle may be adjusted to allow for differences in hand size.

## III. The Apparatus in Use

In making use of the instrument, it is best to allow the prospective subject a practice period once daily for about a week or more before beginning the actual recording. The training period allows the subject to get the 'feel' of the instrument and to learn how to grip in time with a metronome beating once every half second. In addition, this period serves to iron out some of the so-called 'practise effects' of the initial improvement which often crops up in the early course of experimentation. It is surprising to note how difficult it really is for some subjects to learn to grasp in rhythm, so that a practise period is almost essential.

When the recording is to begin, the subject is given the instrument fitted with smoked kymograph paper. Smoking is best accomplished by fastening the paper to the instrument first and then passing the face of the instrument over an ordinary smoking jet. The stylus has been moved to the distal end of the tubular hand and the string is taut enough about the spool to begin instant winding when the alarm clock lever is raised. The subject grasps the dynamometer in his preferred hand allowing it to hang loosely by his side. The metronome is set beating at 120 per minute and the subject signals when he has caught the 'feel' of the rhythm. The experimenter trips the clock mechanism, and the subject grips in time until exhausted. No single trial has been found to last two minutes, and, among the college men and women studied by us, the greatest endurance was a minute and a half—

not long enough to cause the stylus-rod to come to complete rest.

When the recording is complete, the necessary data such as the subject's name, the date, etc., may be scratched in one of the corners with a wire or bamboo stylus. The record is unfastened and run through the usual fixing bath of alcohol and shellac. In order to read the points of such a work-curve, it will be necessary to construct a scale graded in kilograms pulled from the original dial of the Smedley instrument. This scale can be drawn or scratched on transparent celluloid as in Fig. 4, or etched in glass. The recording is placed on a drawing board and the scale is placed over it so that the base line or bottom edge of the scale coincides with the base line of the recording, and the arched section of the scale (*a*, Fig. 4) coincides with the first arc on the recording which is always traced by the rider on the end of the tube. Figure 3 shows this heavy, outer guide-line as it is clearly defined in a sample re-

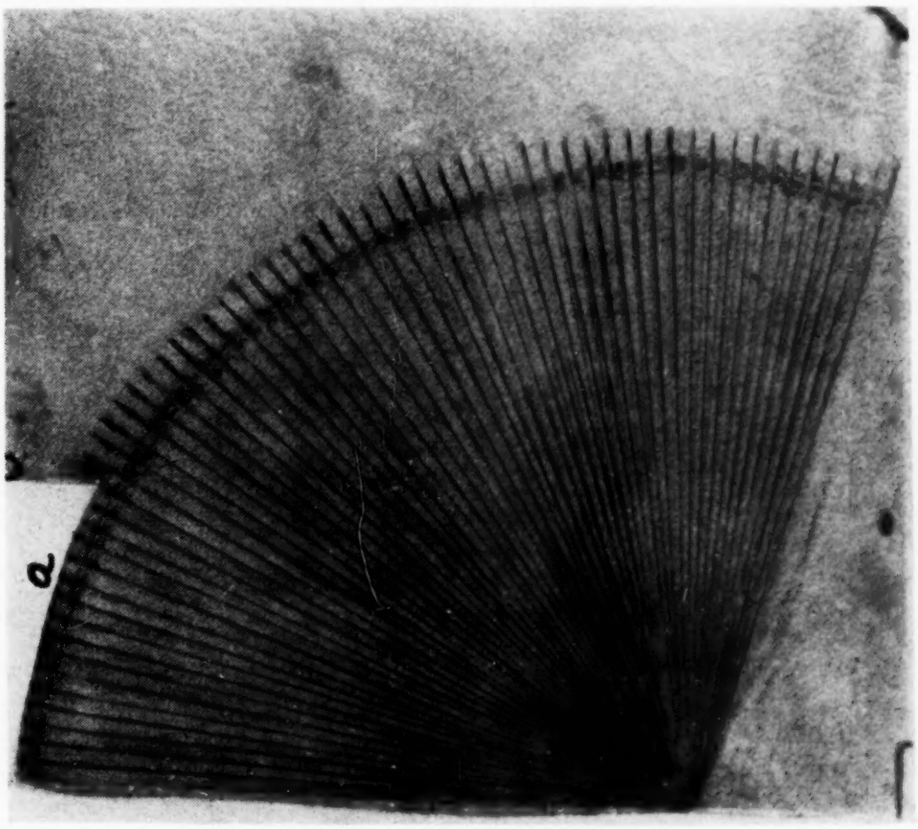


FIG. 4. Scale in kilograms for reading recordings.

cording. The readings taken from the top points of such a recording can be plotted graphically as is convenient to give the usual picture of a fatigue curve. The points along the bottom of the recording can also be read from the scale and plotted to show contracture, or the amount of holding which varies from subject to subject.

## IV. Summary

A critique of past methods in ergography and dynamometry has been presented. An analysis of this sort has led us to the construction of an apparatus designed to correct the errors usually encountered in recording the work decrement by these methods. The instrument is essentially the standard Smedley Revised Hand Dynamometer remodeled to produce a direct and satisfactory recording of the fatigue curve of the muscles involved in successive, forceful gripping. On the basis of careful testing in an experimental study to be reported soon, we feel that the apparatus is a decided improvement over the usual ergographic and dynamometric procedures for the following reasons:

1. Contractions are made against an isometric spring thus eliminating the changes in extent of muscle movement with their concommittant changes in muscular nutritional state such as are encountered in isotonic and oxotonic contractions.

2. A true picture of the work decrement in voluntary muscle effort is rendered possible including such phenomena as contracture, treppe, initial spurt and end spurt when such are present in individual work curves.

3. The mechanical errors which falsify results in ergography and pneumatic recording on the dynamometer are eliminated.

(Manuscript received May 3, 1940)

### References

1. Ash, I. E., Fatigue and its effects upon control, *Arch. Psychol.*, 1914, **4**, 61.
2. Bergstrom, J. A., A new type of ergograph with a discussion of ergographic experimentation, *Amer. J. Psychol.*, 1903, **14**, 246–276.
3. Binet and Vaschide. Cited by Whipple, G. M., *Manual of Mental and Physical Tests*, Baltimore: Warwick and York, 1910, pp. 91–92.

4. CATTELL, J. M., An ergometer, *Science*, 1897, **5**, 909–910.
5. CRAWLEY, S. L., An experimental investigation of recovery from work, *Arch. Psychol.*, 1936, **13**, 66.
6. FRANZ, S. I., Force of voluntary muscular contractions, *Amer. J. Physiol.*, 1901, **4**, 348–372.
7. GLICK, H. N., *Studies on Fatigue*, Unpublished thesis, Northwestern University, 1914.
8. LOMBARD, W. P., The effect of fatigue on voluntary muscular contraction, *Amer. J. Psychol.*, 1892, **3**, 24–42.
9. MANZER, C. W., An experimental study of rest pauses, *Arch. Psychol.*, 1927, **14**, 84.
10. MOSSO, A., *Fatigue* (Translated by Drummond, M., and W. B. Drummond), N. Y.: Putnam, 1904, pp. xiv–334.
11. WEINLAND, J. D., Variability of performance in the curve of work, *Arch. Psychol.*, 1927, **14**, 68.
12. YOCHELSON, S., *Effects of Rest Pauses on Work Decrement*, Ph.D. Thesis, Yale University, 1930.